****

MPSRCH (TM)

****

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

[redacted]n: Wed Aug 16 09:58:24 2000; MasPar time 4.96 Seconds
482.461 Million cell updates/sec
Tabular output not generated.

Title: >US-09-427-873-2
Description: (1-101) from US09427873.pep
Perfect Score: 101
Sequence: 1 LGKFSQTCYNSAIQGSVLTS.....STKINLDDHIANIDGTLKYE 101

Scoring table: TABLE unitprotable
Gap 60

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%
Listing first 1000 summaries

Database: a-geneseq36
1:geneseqp

Statistics: Mean 2.441; Variance 0.631; scale 3.869

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

[redacted]lt No.	Score	% Query Match	Length	DB	ID	Description	Pred. No.
1	101	100.0	101	1	W80552	Antiviral protein, cya	2.72e-148
2	101	100.0	101	1	W67569	N. ellipsosporum cyano	2.72e-148
3	101	100.0	101	1	W06811	Cyanovirin-N.	2.72e-148
4	101	100.0	109	1	W80553	Antiviral protein, cya	2.72e-148
5	101	100.0	109	1	W67570	FLAG epitope-cyanoviri	2.72e-148
6	101	100.0	109	1	W06812	FLAG-cyanovirin-N fusi	2.72e-148
7	8	7.9	299	1	R95970	Nucleic acid recogniti	3.39e-01
8	7	6.9	443	1	W20743	H. pylori cytoplasmic	5.52e+00
9	6	5.9	13	1	W14832	PDGF-1 oncogene protei	7.69e+01
10	6	5.9	48	1	W79087	Human secreted protein	7.69e+01
11	6	5.9	109	1	R25443	PDGF analogue #3.	7.69e+01
12	6	5.9	109	1	R42810	Mature human PDGF-B.	7.69e+01
13	6	5.9	109	1	P81030	Sequence of mature B-c	7.69e+01
14	6	5.9	109	1	R20967	Sequence of B-chain of	7.69e+01
15	6	5.9	109	1	R87515	PDGF mosaic peptide B2	7.69e+01
16	6	5.9	109	1	R25673	PDGF-B.	7.69e+01
17	6	5.9	109	1	R15643	PDGF B-chain polypepti	7.69e+01
18	6	5.9	110	1	R87517	PDGF mosaic peptide B/	7.69e+01
19	6	5.9	110	1	R25445	PDGF analogue #5.	7.69e+01
20	6	5.9	110	1	R25444	PDGF analogue #4.	7.69e+01
21	6	5.9	110	1	R68619	Mature human PDGF-B.	7.69e+01
22	6	5.9	110	1	R87516	PDGF mosaic peptide B/	7.69e+01
23	6	5.9	114	1	P80163	Biosynthetic multifunc	7.69e+01
24	6	5.9	114	1	R04020	PDGF-BB monomer unit.	7.69e+01
25	6	5.9	114	1	R26047	PDGF-B7.	7.69e+01
26	6	5.9	120	1	R41523	rPDGF refolded B119 ho	7.69e+01
27	6	5.9	120	1	R67261	PDGF-B119.	7.69e+01
28	6	5.9	120	1	R12879	rPDGF-B119 from pCFM11	7.69e+01
29	6	5.9	120	1	R60614	Human PDGF-B 119 subun	7.69e+01
30	6	5.9	129	1	R12881	rPDGF B109 [Ser43, Ser	7.69e+01
31	6	5.9	146	1	W64063	Chimeric rhPDGF-B expr	7.69e+01
32	6	5.9	146	1	W64062	Chimeric rhPDGF-B prot	7.69e+01
33	6	5.9	161	1	R26045	Synthetic PDGF-B.	7.69e+01
34	6	5.9	175	1	R87507	Caenorhabditis elegans	7.69e+01
35	6	5.9	190	1	R50012	Truncated Platelet der	7.69e+01
36	6	5.9	190	1	R60615	Human PDGF-B 109 subun	7.69e+01
37	6	5.9	198	1	R05615	BIV vtf gene product.	7.69e+01
38	6	5.9	202	1	W53971	Crinipellis scabella w	7.69e+01
39	6	5.9	204	1	W53970	Myceliophthora thermop	7.69e+01
40	6	5.9	205	1	W64064	Chimeric rhPDGF-B prot	7.69e+01
41	6	5.9	205	1	W64066	Chimeric rhPDGF-B prot	7.69e+01
42	6	5.9	216	1	R68617	Human PDGF-B precursor	7.69e+01
43	6	5.9	220	1	R63470	Recombinant platelet d	7.69e+01
44	6	5.9	220	1	R40965	rPDGF B.	7.69e+01
45	6	5.9	225	1	W04925	Cellulytic enzyme #1 o	7.69e+01
46	6	5.9	226	1	P60215	Sequence encoded by th	7.69e+01
47	6	5.9	226	1	W04930	Cellulytic enzyme #6 o	7.69e+01
48	6	5.9	226	1	P81029	Sequence encoded by th	7.69e+01
49	6	5.9	226	1	R22673	v-sis protein p28sis.	7.69e+01
50	6	5.9	241	1	R63472	Recombinant platelet d	7.69e+01
51	6	5.9	241	1	R50009	Platelet-derived Growt	7.69e+01
52	6	5.9	241	1	R50002	Platelet-derived Growt	7.69e+01
53	6	5.9	241	1	P80596	Recombinant platelet d	7.69e+01
54	6	5.9	241	1	R38919	Human Platelet Derived	7.69e+01
55	6	5.9	241	1	R40967	CV-sis gene product.	7.69e+01
56	6	5.9	241	1	W64065	Chimeric rhPDGF-B prot	7.69e+01
57	6	5.9	241	1	R40964	PDGF Bc-sis.	7.69e+01
58	6	5.9	241	1	P80597	Cv-sis gene encoded Pl	7.69e+01
59	6	5.9	241	1	R63469	Recombinant platelet d	7.69e+01
60	6	5.9	261	1	W31729	Mycobacterium tubercul	7.69e+01
61	6	5.9	264	1	W89843	Protein encoded by clo	7.69e+01
62	6	5.9	265	1	W89920	Antigen 5 from cluster	7.69e+01
63	6	5.9	271	1	R40963	PDGF Bv-sis.	7.69e+01
64	6	5.9	271	1	P80595	Recombinant platelet d	7.69e+01
65	6	5.9	271	1	R63468	Recombinant platelet d	7.69e+01
66	6	5.9	282	1	R60616	Human PDGF-B 119 linke	7.69e+01
67	6	5.9	293	1	W04936	Chimeric endoglucanase	7.69e+01
68	6	5.9	297	1	W04933	Chimeric endoglucanase	7.69e+01
69	6	5.9	298	1	W04935	Chimeric endoglucanase	7.69e+01
70	6	5.9	307	1	W81823	Human VTAP-2 protein.	7.69e+01
71	6	5.9	308	1	W04934	Chimeric endoglucanase	7.69e+01
72	6	5.9	348	1	W27267	Streptococcus pneumoni	7.69e+01
73	6	5.9	355	1	R98154	Blowfly PM48 antigen.	7.69e+01
74	6	5.9	374	1	W79743	Crepis sp. delta-12-ep	7.69e+01
75	6	5.9	374	1	W79742	Crepis palaestina delt	7.69e+01
76	6	5.9	376	1	R38267	Sequence encoded by pa	7.69e+01
77	6	5.9	402	1	R05277	New serine hydroxymeth	7.69e+01
78	6	5.9	404	1	R78621	Chicken GalNAc-alpha-2	7.69e+01
79	6	5.9	434	1	R97745	Serine hydroxymethyl t	7.69e+01
80	6	5.9	450	1	Y00170	Enterococcus faecalis	7.69e+01
81	6	5.9	485	1	Y00895	Rat GMEB-2' protein se	7.69e+01
82	6	5.9	510	1	W04334	Human lymphoid cell ac	7.69e+01
83	6	5.9	510	1	W04264	Human CD39 protein.	7.69e+01
84	6	5.9	528	1	Y00031	Enterococcus faecalis	7.69e+01
85	6	5.9	529	1	Y00894	Rat GMEB-2 protein seq	7.69e+01
86	6	5.9	547	1	Y00030	Enterococcus faecalis	7.69e+01
87	6	5.9	549	1	R92050	KM31-7 precursor.	7.69e+01
88	6	5.9	551	1	W83404	Human KM-102-derived r	7.69e+01
89	6	5.9	626	1	W41501	Human DP.75, a putativ	7.69e+01
90	6	5.9	678	1	R42087	D. melanogaster dorsal	7.69e+01
91	6	5.9	686	1	P61518	Sequence of cyclomalto	7.69e+01
92	6	5.9	739	1	R38266	Sequence of a bovine p	7.69e+01
93	6	5.9	743	1	W03319	Tobacco transketolase.	7.69e+01
94	6	5.9	765	1	W89914	Antigen from cluster	7.69e+01
95	6	5.9	911	1	R52568	PEPCase gene in lambda	7.69e+01
96	6	5.9	912	1	W42103	Human p115 Rho-guanine	7.69e+01

97	6	5.9	1266	1	W48721	Human cytoplasmic isol	7.69e+01
98	6	5.9	1475	1	R08221	Recombinant alpha amyl	7.69e+01
99	5	5.0	12	1	R63585	Conserved region 2 fro	8.72e+02
100	5	5.0	12	1	R90560	Human influenza A viru	8.72e+02
101	5	5.0	14	1	R76589	Control peptide from P	8.72e+02
102	5	5.0	16	1	R22457	Peptide encoded by DNA	8.72e+02
103	5	5.0	17	1	R79941	Jojoba wax-synthase pe	8.72e+02
104	5	5.0	17	1	R77181	Jojoba 57 kDa wax synt	8.72e+02
105	5	5.0	17	1	W79718	gp120 immunogenic pept	8.72e+02
106	5	5.0	20	1	R56619	Alginic acid Al-II-2 l	8.72e+02
107	5	5.0	20	1	R77618	Ascorbate oxidase N-te	8.72e+02
108	5	5.0	20	1	R61238	Alginate lyase II N-te	8.72e+02
109	5	5.0	21	1	Y11611	Human 5' EST secreted	8.72e+02
110	5	5.0	21	1	R22455	BGP N-terminal fractio	8.72e+02
111	5	5.0	21	1	R46479	Consensus sequence of	8.72e+02
112	5	5.0	22	1	W79434	Staphylococcus aureus	8.72e+02
113	5	5.0	22	1	R81711	C.albicans enolase pep	8.72e+02
114	5	5.0	26	1	W79433	Staphylococcus aureus	8.72e+02
115	5	5.0	27	1	W89019	Sequence ID #719 from	8.72e+02
116	5	5.0	29	1	R75093	Rat SIII 15 kDa subuni	8.72e+02
117	5	5.0	30	1	P80653	Peptide encoded by pro	8.72e+02
118	5	5.0	30	1	R05809	Signal peptide derived	8.72e+02
119	5	5.0	37	1	R72967	Pig kidney cell mutaro	8.72e+02
120	5	5.0	39	1	R60167	Rice mitochondrial pro	8.72e+02
121	5	5.0	42	1	W62694	Streptococcus pneumoni	8.72e+02
122	5	5.0	44	1	W48346	Human breast cancer re	8.72e+02
123	5	5.0	44	1	W88624	Secreted protein encod	8.72e+02
124	5	5.0	44	1	R51626	p53 antisense strand p	8.72e+02
125	5	5.0	47	1	W26679	5.0 kDa bacteriocalcif	8.72e+02
126	5	5.0	50	1	W26675	Bacteriocalcifin-1 fro	8.72e+02
127	5	5.0	50	1	W45409	N-terminal amino acid	8.72e+02
128	5	5.0	51	1	W32339	Mycobacterium tubercul	8.72e+02
129	5	5.0	51	1	W64334	Mycobacterium tubercul	8.72e+02
130	5	5.0	51	1	W32466	Mycobacterium tubercul	8.72e+02
131	5	5.0	51	1	W81701	M. tuberculosis immuno	8.72e+02
132	5	5.0	55	1	W99675	Enzyme component conta	8.72e+02
133	5	5.0	55	1	Y04113	Enzyme component #2.	8.72e+02
134	5	5.0	56	1	Y11192	S. pneumoniae L-lactat	8.72e+02
135	5	5.0	56	1	R64243	CCSL based on HPV-6 DN	8.72e+02
136	5	5.0	66	1	W13014	MF2 protein induces pa	8.72e+02
137	5	5.0	69	1	R39108	LD78 Gln21>Ser.	8.72e+02
138	5	5.0	69	1	R09348	Sequence corresp. to r	8.72e+02
139	5	5.0	69	1	W46470	Amino acid sequence of	8.72e+02
140	5	5.0	69	1	R03105	CD4 receptor protein a	8.72e+02
141	5	5.0	69	1	W27906	Amino acid sequence of	8.72e+02
142	5	5.0	70	1	W85636	Antigenic epitope of Z	8.72e+02
143	5	5.0	77	1	Y12267	Human 5' EST secreted	8.72e+02
144	5	5.0	78	1	W88658	Secreted protein encod	8.72e+02
145	5	5.0	82	1	W61202	Streptococcus pneumoni	8.72e+02
146	5	5.0	89	1	W98215	H. pylori GHPO 1213 pr	8.72e+02
147	5	5.0	96	1	R70799	Human secreted protein	8.72e+02
148	5	5.0	100	1	W85630	Antigenic epitope of Z	8.72e+02
149	5	5.0	101	1	R47067	HaEPV p11.5 ORF.	8.72e+02
150	5	5.0	101	1	W15397	HaEPV 11 kDa insectici	8.72e+02
151	5	5.0	102	1	Y11125	S. pneumoniae protein	8.72e+02
152	5	5.0	103	1	Y00037	Enterococcus faecalis	8.72e+02
153	5	5.0	104	1	R52045	Heavy chain variable r	8.72e+02
154	5	5.0	109	1	W22724	Sinapis alba flowering	8.72e+02
155	5	5.0	109	1	W22486	Phaffia derived glycer	8.72e+02
156	5	5.0	110	1	W22725	Arabidopsis flowering	8.72e+02
157	5	5.0	111	1	R40173	Humanised IL-5 antibod	8.72e+02
158	5	5.0	113	1	W58486	Murine HYH antibody he	8.72e+02
159	5	5.0	113	1	R38608	HYH heavy chain.	8.72e+02
160	5	5.0	122	1	R52663	FNB curlin.	8.72e+02
161	5	5.0	124	1	W38583	Streptococcus pneumoni	8.72e+02
162	5	5.0	124	1	W56091	Human secreted protein	8.72e+02
163	5	5.0	125	1	R44591	Monomeric PDGF-A.	8.72e+02
164	5	5.0	126	1	W97351	Amino acid sequence of	8.72e+02
165	5	5.0	126	1	Y11047	H. pylori ORF hp2e1091	8.72e+02
166	5	5.0	127	1	W31265	Neospora caninum antig	8.72e+02
167	5	5.0	128	1	W97716	Staphylococcus aureus	8.72e+02
168	5	5.0	128	1	Y00036	Enterococcus faecalis	8.72e+02
169	5	5.0	128	1	W85643	Human Zcyto7 mature pr	8.72e+02
170	5	5.0	135	1	R40175	Humanised antibody CMX	8.72e+02
171	5	5.0	135	1	R40179	Humanised antibody CMX	8.72e+02
172	5	5.0	135	1	R40177	Humanised antibody CMX	8.72e+02
173	5	5.0	135	1	R40183	Humanised antibody CMX	8.72e+02
174	5	5.0	140	1	R87058	CDR grafted anti-IL-5	8.72e+02
175	5	5.0	140	1	W88981	Polypeptide fragment e	8.72e+02
176	5	5.0	141	1	R70979	Crocodile alpha-globin	8.72e+02
177	5	5.0	142	1	R52664	Fibronectin bind prote	8.72e+02
178	5	5.0	143	1	R41533	M.fermentans incognitu	8.72e+02
179	5	5.0	144	1	W28237	Staphylococcus aureus	8.72e+02
180	5	5.0	146	1	W73397	Human secreted protein	8.72e+02
181	5	5.0	148	1	Y12370	Human 5' EST secreted	8.72e+02
182	5	5.0	151	1	W85641	Human Zcyto7 mature pr	8.72e+02
183	5	5.0	152	1	W20280	H. pylori cytoplasmic	8.72e+02
184	5	5.0	153	1	W85642	Human Zcyto7 mature pr	8.72e+02
185	5	5.0	154	1	W85640	Human Zcyto7 mature pr	8.72e+02
186	5	5.0	156	1	W69759	Acetobacter xylinum bc	8.72e+02
187	5	5.0	156	1	R45003	Cellulose synthase ope	8.72e+02
188	5	5.0	157	1	R81709	Recombinant C.albicans	8.72e+02
189	5	5.0	157	1	W85644	Human Zcyto7 mature pr	8.72e+02
190	5	5.0	158	1	W85639	Human Zcyto7 mature pr	8.72e+02
191	5	5.0	159	1	W18654	Human microtubule-asso	8.72e+02
192	5	5.0	160	1	W85618	Human Zcyto7 mature pr	8.72e+02
193	5	5.0	160	1	W85625	Human Zcyto7 mature pr	8.72e+02
194	5	5.0	160	1	W85619	Human Zcyto7 mature pr	8.72e+02
195	5	5.0	160	1	W85629	Human Zcyto7 mature pr	8.72e+02
196	5	5.0	160	1	W85624	Human Zcyto7 mature pr	8.72e+02
197	5	5.0	160	1	W85623	Human Zcyto7 mature pr	8.72e+02
198	5	5.0	160	1	W85620	Human Zcyto7 mature pr	8.72e+02
199	5	5.0	160	1	W85627	Human Zcyto7 mature pr	8.72e+02
200	5	5.0	160	1	W85621	Human Zcyto7 mature pr	8.72e+02
201	5	5.0	160	1	W85626	Human Zcyto7 mature pr	8.72e+02
202	5	5.0	160	1	W85622	Human Zcyto7 mature pr	8.72e+02
203	5	5.0	160	1	W85628	Human Zcyto7 mature pr	8.72e+02
204	5	5.0	163	1	R60215	Immunogenic fragment o	8.72e+02
205	5	5.0	163	1	R38877	Sequence of haemagglut	8.72e+02
206	5	5.0	165	1	Y04934	Mycobacterium species	8.72e+02
207	5	5.0	166	1	W67636	MSRV-1 virus clone LB1	8.72e+02
208	5	5.0	166	1	W85048	Amino acid sequence of	8.72e+02
209	5	5.0	169	1	W42106	Amino acid sequence of	8.72e+02
210	5	5.0	180	1	W97350	Interleukin-20.	8.72e+02
211	5	5.0	180	1	W85615	Human Zcyto7.	8.72e+02
212	5	5.0	185	1	W41975	Flea serine protease S	8.72e+02
213	5	5.0	188	1	R70744	PagC protein.	8.72e+02
214	5	5.0	188	1	W18380	S. typhimurium pagC ge	8.72e+02
215	5	5.0	188	1	R26415	pagC/AP fusion protein	8.72e+02
216	5	5.0	191	1	W94495	Human delta-2 protein	8.72e+02
217	5	5.0	191	1	W73377	Human HPDDV78 protein	8.72e+02
218	5	5.0	194	1	R99852	Proteinase A subunit.	8.72e+02
219	5	5.0	198	1	W72394	Pathogen response prot	8.72e+02
220	5	5.0	208	1	W75130	Human secreted protein	8.72e+02
221	5	5.0	212	1	W98872	H. pylori GHPO 1732 pr	8.72e+02
222	5	5.0	213	1	W99164	RasGAP catalytic domai	8.72e+02
223	5	5.0	213	1	W99179	RasGAP catalytic domai	8.72e+02
224	5	5.0	216	1	W68412	Hybrid alpha-1-thymosi	8.72e+02
225	5	5.0	218	1	W30839	MyokL protein myokLp.	8.72e+02
226	5	5.0	218	1	W15104	Myxoma virus MA56 gene	8.72e+02
227	5	5.0	220	1	Y11027	H. pylori ORF 02ge4162	8.72e+02
228	5	5.0	220	1	R88272	Papilloma virus major	8.72e+02
229	5	5.0	221	1	R88273	Papilloma virus major	8.72e+02
230	5	5.0	222	1	R38866	Sequence of the HA2 su	8.72e+02
231	5	5.0	222	1	R60222	Influenza haemagglutin	8.72e+02
232	5	5.0	223	1	W97837	Tobacco TTG1 homologue	8.72e+02
233	5	5.0	224	1	R88271	Papilloma virus major	8.72e+02
234	5	5.0	226	1	W72902	Mycobacterium tubercul	8.72e+02
235	5	5.0	229	1	W49694	Human ScFv5 against al	8.72e+02
236	5	5.0	236	1	W49690	Human ScFv1 against al	8.72e+02
237	5	5.0	237	1	W49691	Human ScFv2 against al	8.72e+02
238	5	5.0	237	1	W23085	Microscilla furvescens	8.72e+02
239	5	5.0	239	1	W49692	Human ScFv3 against al	8.72e+02
240	5	5.0	240	1	R44027	Lys-63 cholera toxin s	8.72e+02
241	5	5.0	244	1	W68593	Tiarosporella phaseoli	8.72e+02
242	5	5.0	249	1	W24061	Human WSX receptor ago	8.72e+02

243	5	5.0	249	1	R34549	Putative calcium chann	8.72e+02
244	5	5.0	251	1	R97012	Influenza virus matrix	8.72e+02
245	5	5.0	252	1	R32021	Influenza matrix prote	8.72e+02
246	5	5.0	252	1	R36804	Influenza A virus matr	8.72e+02
247	5	5.0	254	1	W49693	Human ScFv4 against al	8.72e+02
248	5	5.0	254	1	P70263	Protein comprising sig	8.72e+02
249	5	5.0	256	1	W75103	Human secreted protein	8.72e+02
250	5	5.0	257	1	W55602	H. pylori ORF 11ap2071	8.72e+02
251	5	5.0	259	1	R94586	C. pneumoniae polypept	8.72e+02
252	5	5.0	259	1	W01743	C. pneumoniae 53 kDa a	8.72e+02
253	5	5.0	263	1	R22957	Human proteasome compo	8.72e+02
254	5	5.0	263	1	W69232	FcR-II protein sequenc	8.72e+02
255	5	5.0	265	1	W74905	Human secreted protein	8.72e+02
256	5	5.0	265	1	R80063	Human IFNAB-BPI encode	8.72e+02
257	5	5.0	269	1	W98462	H. pylori GHPO 726 pro	8.72e+02
258	5	5.0	269	1	R60210	Immunogenic fragment o	8.72e+02
259	5	5.0	269	1	R38872	Sequence of H1 subtype	8.72e+02
260	5	5.0	269	1	R22666	Protein used to raise	8.72e+02
261	5	5.0	269	1	W71545	Helicobacter polypepti	8.72e+02
262	5	5.0	271	1	R94580	C. pneumoniae polypept	8.72e+02
263	5	5.0	273	1	W77635	Potassium-transporting	8.72e+02
264	5	5.0	275	1	W55565	H. pylori ORF 06cp112i	8.72e+02
265	5	5.0	278	1	W55270	H. pylori ORF 09ap9080	8.72e+02
266	5	5.0	279	1	R94526	Korean Viper Salmosa t	8.72e+02
267	5	5.0	281	1	R25719	Prod. of gene regulati	8.72e+02
268	5	5.0	282	1	R22752	Proctase-A.	8.72e+02
269	5	5.0	283	1	W36025	Fragment of MSRV-1 gag	8.72e+02
270	5	5.0	283	1	W37469	Connexin-32.	8.72e+02
271	5	5.0	283	1	W71085	Multiple sclerosis ass	8.72e+02
272	5	5.0	284	1	W13238	Human steroidogenesis	8.72e+02
273	5	5.0	285	1	W13237	Human steroidogenesis	8.72e+02
274	5	5.0	285	1	W13236	Human steroidogenesis	8.72e+02
275	5	5.0	285	1	W00814	Human steroidogenesis	8.72e+02
276	5	5.0	285	1	W13234	Human steroidogenesis	8.72e+02
277	5	5.0	286	1	R88917	Modified IncC protein,	8.72e+02
278	5	5.0	288	1	Y05394	Human TIE ligand NL3 p	8.72e+02
279	5	5.0	289	1	W48804	Homo sapiens clone BK1	8.72e+02
280	5	5.0	289	1	W01128	Coat protein (short ve	8.72e+02
281	5	5.0	292	1	R04895	Penicillinase-Insulin	8.72e+02
282	5	5.0	292	1	W61001	Streptococcus pneumoni	8.72e+02
283	5	5.0	292	1	R90545	pJG4-5-CDK-BP clone #1	8.72e+02
284	5	5.0	293	1	W44873	Murine BIN-1 Associate	8.72e+02
285	5	5.0	299	1	W89198	Aspartate-specific cys	8.72e+02
286	5	5.0	301	1	R12315	Peptide fragment of HI	8.72e+02
287	5	5.0	303	1	R46266	Improved Heat-stable c	8.72e+02
288	5	5.0	303	1	R46264	Improved Heat-stable c	8.72e+02
289	5	5.0	303	1	R46253	Heat-stable carbamylas	8.72e+02
290	5	5.0	303	1	R46240	Heat-stable carbamylas	8.72e+02
291	5	5.0	303	1	R46262	Improved Heat-stable c	8.72e+02
292	5	5.0	303	1	R46259	Improved Heat-stable c	8.72e+02
293	5	5.0	303	1	R46242	Heat-stable carbamylas	8.72e+02
294	5	5.0	303	1	R46244	Heat-stable carbamylas	8.72e+02
295	5	5.0	303	1	R46263	Improved Heat-stable c	8.72e+02
296	5	5.0	303	1	R46256	Improved Heat-stable c	8.72e+02
297	5	5.0	303	1	R46247	Heat-stable carbamylas	8.72e+02
298	5	5.0	303	1	R46260	Improved Heat-stable c	8.72e+02
299	5	5.0	303	1	R46241	Heat-stable carbamylas	8.72e+02
300	5	5.0	303	1	R46272	Improved Heat-stable c	8.72e+02
301	5	5.0	303	1	R46270	Improved Heat-stable c	8.72e+02
302	5	5.0	303	1	R46269	Improved Heat-stable c	8.72e+02
303	5	5.0	303	1	R46258	Improved Heat-stable c	8.72e+02
304	5	5.0	303	1	R46265	Improved Heat-stable c	8.72e+02
305	5	5.0	303	1	R46243	Heat-stable carbamylas	8.72e+02
306	5	5.0	303	1	R46245	Heat-stable carbamylas	8.72e+02
307	5	5.0	303	1	R46251	Heat-stable carbamylas	8.72e+02
308	5	5.0	303	1	R46267	Improved Heat-stable c	8.72e+02
309	5	5.0	303	1	R46255	Improved Heat-stable c	8.72e+02
310	5	5.0	303	1	R46252	Heat-stable carbamylas	8.72e+02
311	5	5.0	303	1	R46249	Heat-stable carbamylas	8.72e+02
312	5	5.0	303	1	R25358	KNK-712.	8.72e+02
313	5	5.0	303	1	R46257	Improved Heat-stable c	8.72e+02
314	5	5.0	303	1	R46268	Improved Heat-stable c	8.72e+02
315	5	5.0	303	1	R46246	Heat-stable carbamylas	8.72e+02
316	5	5.0	303	1	R46248	Heat-stable carbamylas	8.72e+02
317	5	5.0	303	1	R46261	Improved Heat-stable c	8.72e+02
318	5	5.0	303	1	R46239	Heat-stable carbamylas	8.72e+02
319	5	5.0	304	1	W70445	Mutant D-N-alpha-carba	8.72e+02
320	5	5.0	304	1	W13405	Eupenicillium brefeldi	8.72e+02
321	5	5.0	304	1	R82836	Agrobacterium radiobac	8.72e+02
322	5	5.0	304	1	R38880	Sequence of haemagglut	8.72e+02
323	5	5.0	304	1	R53639	Carbamoylase enzyme of	8.72e+02
324	5	5.0	304	1	W70448	Mutant D-N-alpha-carba	8.72e+02
325	5	5.0	304	1	W70449	Mutant D-N-alpha-carba	8.72e+02
326	5	5.0	304	1	W70446	Mutant D-N-alpha-carba	8.72e+02
327	5	5.0	304	1	W70447	Mutant D-N-alpha-carba	8.72e+02
328	5	5.0	305	1	R93612	Kaposi's sarcoma assoc	8.72e+02
329	5	5.0	305	1	R97836	Kaposi's sarcoma assoc	8.72e+02
330	5	5.0	306	1	W80671	S. pneumoniae protein	8.72e+02
331	5	5.0	306	1	W60856	FabD polypeptide seque	8.72e+02
332	5	5.0	307	1	R38870	Sequence of C13 which	8.72e+02
333	5	5.0	307	1	R04207	C13 protein - a recomb	8.72e+02
334	5	5.0	309	1	W83322	Single chain Apo-2 ant	8.72e+02
335	5	5.0	309	1	W98581	H. pylori GHPO 65 prot	8.72e+02
336	5	5.0	311	1	W53549	Human Cdx2 protein.	8.72e+02
337	5	5.0	312	1	Y05528	Soybean isoflavone red	8.72e+02
338	5	5.0	312	1	W83323	Single chain Apo-2 ant	8.72e+02
339	5	5.0	312	1	W25776	JTV1 protein.	8.72e+02
340	5	5.0	314	1	R34763	OMT1.A and OMT1.B tran	8.72e+02
341	5	5.0	315	1	W53922	Decaprenyl diphosphate	8.72e+02
342	5	5.0	315	1	W19946	Alzheimer's disease pr	8.72e+02
343	5	5.0	317	1	R60599	Hornet phospholipase D	8.72e+02
344	5	5.0	317	1	P90500	Pseudomonas cepacia li	8.72e+02
345	5	5.0	323	1	W05832	M. tuberculosis RNA po	8.72e+02
346	5	5.0	324	1	W25117	CD2-associated intrace	8.72e+02
347	5	5.0	324	1	W80421	CD2 associated intrace	8.72e+02
348	5	5.0	324	1	W26497	CD2 associated intrace	8.72e+02
349	5	5.0	325	1	W82241	A protein with stereos	8.72e+02
350	5	5.0	325	1	W00504	N.meningitidis IM2394	8.72e+02
351	5	5.0	328	1	W55486	H. pylori ORF hp2p1027	8.72e+02
352	5	5.0	329	1	W87563	A protein with stereos	8.72e+02
353	5	5.0	330	1	W98878	H. pylori GHPO 1751 pr	8.72e+02
354	5	5.0	330	1	W98788	H. pylori GHPO 1223 pr	8.72e+02
355	5	5.0	330	1	W87564	A protein with stereos	8.72e+02
356	5	5.0	331	1	R80066	Human IFNAB-BPI.	8.72e+02
357	5	5.0	334	1	W19854	Human purinergic recep	8.72e+02
358	5	5.0	334	1	W22732	Human ATP receptor.	8.72e+02
359	5	5.0	335	1	W82830	Endothelial MAD intera	8.72e+02
360	5	5.0	335	1	W75857	Human secretory protei	8.72e+02
361	5	5.0	336	1	W23397	TWIK-1 potassium chann	8.72e+02
362	5	5.0	336	1	R97986	F22b7.7 potassium chan	8.72e+02
363	5	5.0	338	1	W55522	H. pylori ORF 14ap1022	8.72e+02
364	5	5.0	339	1	W99578	C.albicans Cdk activat	8.72e+02
365	5	5.0	340	1	R98887	N.meningitidis IM2394	8.72e+02
366	5	5.0	340	1	W14644	N. meningitidis IM2394	8.72e+02
367	5	5.0	341	1	R70142	Porcine mutarotase (MU	8.72e+02
368	5	5.0	342	1	W36040	Alfalfa cinnamoyl CoA	8.72e+02
369	5	5.0	344	1	Y02670	Human secreted protein	8.72e+02
370	5	5.0	344	1	R10864	Lipase modulating fact	8.72e+02
371	5	5.0	344	1	R77246	P. cepacia lipase acce	8.72e+02
372	5	5.0	344	1	R39397	Pseudomonas cepacia DS	8.72e+02
373	5	5.0	344	1	R03989	Lambda-SM70 encoding M	8.72e+02
374	5	5.0	346	1	R27576	ABN-A from A. niger.	8.72e+02
375	5	5.0	348	1	W53920	Decaprenyl diphosphate	8.72e+02
376	5	5.0	350	1	W98479	H. pylori GHPO 946 pro	8.72e+02
377	5	5.0	350	1	R63143	Glycoprotein 63 (gp63)	8.72e+02
378	5	5.0	351	1	R41542	B15R product.	8.72e+02
379	5	5.0	351	1	R24251	Vaccinia virus B18R pr	8.72e+02
380	5	5.0	352	1	W99556	Protein encoded by clo	8.72e+02
381	5	5.0	352	1	W55594	H. pylori ORF 09cp6100	8.72e+02
382	5	5.0	352	1	W71070	Multiple sclerosis ass	8.72e+02
383	5	5.0	354	1	W94656	Xenopus WA545 protein.	8.72e+02
384	5	5.0	355	1	Y11081	H. pylori ORF hp1p1393	8.72e+02
385	5	5.0	355	1	W20563	H. pylori cell envelop	8.72e+02
386	5	5.0	357	1	R56479	Sorbitol dehydrogenase	8.72e+02
387	5	5.0	357	1	W55299	H. pylori ORF 14ge1070	8.72e+02
388	5	5.0	357	1	R33440	Ornithine cyclodeamina	8.72e+02

389	5	5.0	358	1	W20735	H. pylori cell envelop	8.72e+02
390	5	5.0	359	1	W86316	Kidney injury associat	8.72e+02
391	5	5.0	360	1	W20878	H. pylori cytoplasmic	8.72e+02
392	5	5.0	364	1	W88979	Polypeptide fragment e	8.72e+02
393	5	5.0	364	1	R34764	OMTI translation produ	8.72e+02
394	5	5.0	365	1	R12316	HIV-2 env peptide frag	8.72e+02
395	5	5.0	366	1	W55624	H. pylori ORF hp6p1090	8.72e+02
396	5	5.0	366	1	W55625	H. pylori ORF hp6p1090	8.72e+02
397	5	5.0	367	1	R54666	Outer membrane lipopro	8.72e+02
398	5	5.0	367	1	W98294	H. pylori GHPO 1694 pr	8.72e+02
399	5	5.0	367	1	R98899	Coat protein (long ver	8.72e+02
400	5	5.0	370	1	W55253	H. pylori ORF 05gp1190	8.72e+02
401	5	5.0	371	1	W55699	H. pylori ORF 06ap1060	8.72e+02
402	5	5.0	373	1	Y04940	Mycobacterium species	8.72e+02
403	5	5.0	375	1	R26327	Asymmetrically active	8.72e+02
404	5	5.0	377	1	W75063	Human secreted protein	8.72e+02
405	5	5.0	378	1	W99558	Protein encoded by pET	8.72e+02
406	5	5.0	378	1	R70205	Alpha-galactosidase of	8.72e+02
407	5	5.0	378	1	W00621	Coffee bean alpha-gala	8.72e+02
408	5	5.0	378	1	W41508	Aplysia cAMP-response	8.72e+02
409	5	5.0	378	1	W71072	Multiple sclerosis ass	8.72e+02
410	5	5.0	378	1	W60766	Murine Lunatic Fringe	8.72e+02
411	5	5.0	378	1	P70542	Sequence encoded by ne	8.72e+02
412	5	5.0	378	1	R37591	Sequence of microsomal	8.72e+02
413	5	5.0	381	1	R69623	S. avermitilis BCKDH E	8.72e+02
414	5	5.0	383	1	W20913	H. pylori secreted or	8.72e+02
415	5	5.0	385	1	W71490	Helicobacter polypepti	8.72e+02
416	5	5.0	388	1	W62298	Wheat D1 protease.	8.72e+02
417	5	5.0	389	1	W00623	SacI methylase.	8.72e+02
418	5	5.0	390	1	Y05302	S. aureus protein sequ	8.72e+02
419	5	5.0	391	1	W37992	Mutant Aspergillus ory	8.72e+02
420	5	5.0	392	1	W31705	Human extracellular/ep	8.72e+02
421	5	5.0	395	1	W71510	Helicobacter polypepti	8.72e+02
422	5	5.0	396	1	W97742	Corn S-adenosylmethion	8.72e+02
423	5	5.0	396	1	W37377	Hepatitis C virus chim	8.72e+02
424	5	5.0	398	1	W99557	Protein encoded by pET	8.72e+02
425	5	5.0	398	1	W98515	H. pylori GHPO 1301 pr	8.72e+02
426	5	5.0	398	1	W71547	Helicobacter polypepti	8.72e+02
427	5	5.0	398	1	W71071	Multiple sclerosis ass	8.72e+02
428	5	5.0	399	1	W05523	HCMV Towne strain UL15	8.72e+02
429	5	5.0	400	1	R56281	Chitin-deacetylase.	8.72e+02
430	5	5.0	400	1	W61233	Streptococcus pneumoni	8.72e+02
431	5	5.0	403	1	R77502	Wild type creatinase.	8.72e+02
432	5	5.0	403	1	W72943	Mycobacterium tubercul	8.72e+02
433	5	5.0	403	1	P60274	Sequence of creatine a	8.72e+02
434	5	5.0	403	1	P80680	Creatine amidohydrolas	8.72e+02
435	5	5.0	404	1	W54077	LH-2 protein #1.	8.72e+02
436	5	5.0	404	1	W72942	Mycobacterium tubercul	8.72e+02
437	5	5.0	404	1	R37594	Sequence of plastid de	8.72e+02
438	5	5.0	406	1	R34550	Putative calcium chann	8.72e+02
439	5	5.0	407	1	R34551	Putative calcium chann	8.72e+02
440	5	5.0	410	1	W53921	Decaprenyl diphosphate	8.72e+02
441	5	5.0	411	1	W80133	Arthrobacter sp. creat	8.72e+02
442	5	5.0	411	1	R70206	Alpha-galactosidase of	8.72e+02
443	5	5.0	412	1	Y11085	H. pylori ORF hp5p1557	8.72e+02
444	5	5.0	412	1	W97051	Protein encoded by ope	8.72e+02
445	5	5.0	412	1	W89246	Sphingomonas sp. strai	8.72e+02
446	5	5.0	414	1	R93982	Saccharothrix aerocolo	8.72e+02
447	5	5.0	416	1	W98286	H. pylori GHPO 1368 pr	8.72e+02
448	5	5.0	416	1	W55613	H. pylori ORF hp4e5339	8.72e+02
449	5	5.0	420	1	R61327	Tau-protein kinase I (	8.72e+02
450	5	5.0	420	1	R22215	Sequence of human inte	8.72e+02
451	5	5.0	420	1	R71326	Coffee bean alpha-gala	8.72e+02
452	5	5.0	420	1	R61326	Tau-protein kinase I (	8.72e+02
453	5	5.0	420	1	W82842	Human interleukin-5 re	8.72e+02
454	5	5.0	420	1	R22219	Sequence of secretory	8.72e+02
455	5	5.0	421	1	R25064	Human IL-5 receptor al	8.72e+02
456	5	5.0	421	1	W37376	Hepatitis C virus chim	8.72e+02
457	5	5.0	422	1	W71639	Omega-cyclohexane fatt	8.72e+02
458	5	5.0	423	1	W54078	LH-2 protein #2.	8.72e+02
459	5	5.0	424	1	W77311	Phaffia rhodozyma endo	8.72e+02
460	5	5.0	424	1	W97851	Hypersensitive respons	8.72e+02
461	5	5.0	427	1	Y04941	Mycobacterium species	8.72e+02
462	5	5.0	428	1	W87975	Rio Mamore Hantavirus	8.72e+02
463	5	5.0	430	1	P70312	Sequence encoded by pr	8.72e+02
464	5	5.0	430	1	R24438	Sequence encoded by pr	8.72e+02
465	5	5.0	430	1	P70224	Sequence of Mucor mieh	8.72e+02
466	5	5.0	430	1	R62932	Mucor miehei carboxyl	8.72e+02
467	5	5.0	432	1	R94585	DHFR/C. pneumoniae ant	8.72e+02
468	5	5.0	434	1	R57434	Serine hydroxymethyl t	8.72e+02
469	5	5.0	435	1	W11230	S-phase kinase associa	8.72e+02
470	5	5.0	435	1	R60500	Linoleic-acid-desatura	8.72e+02
471	5	5.0	436	1	W78193	Human secreted protein	8.72e+02
472	5	5.0	437	1	R06281	Swine enzootic pneumon	8.72e+02
473	5	5.0	440	1	R43658	Rhamnogalacturonase.	8.72e+02
474	5	5.0	440	1	R81708	Recombinant C.albicans	8.72e+02
475	5	5.0	440	1	R11515	Soybean chlorotic mott	8.72e+02
476	5	5.0	441	1	R70209	A. niger alpha-galacto	8.72e+02
477	5	5.0	442	1	W00505	N.meningitidis IM2394	8.72e+02
478	5	5.0	444	1	W36845	Single chain T-cell re	8.72e+02
479	5	5.0	446	1	R37593	Sequence of a plastid	8.72e+02
480	5	5.0	446	1	R60499	Linoleic-acid-desatura	8.72e+02
481	5	5.0	446	1	W04725	Aromatic acyl transfer	8.72e+02
482	5	5.0	447	1	W96260	Hypersensitive respons	8.72e+02
483	5	5.0	447	1	R48200	A.chrysogenum beta-tub	8.72e+02
484	5	5.0	447	1	W61228	Streptococcus pneumoni	8.72e+02
485	5	5.0	447	1	R40226	Acremonium chrysogenum	8.72e+02
486	5	5.0	448	1	W95709	Homo sapiens fetal kid	8.72e+02
487	5	5.0	448	1	W94281	Human extracellular ma	8.72e+02
488	5	5.0	448	1	R56512	Beta tubulin of Acremo	8.72e+02
489	5	5.0	448	1	W79739	Human EEGF protein.	8.72e+02
490	5	5.0	450	1	R96021	P. gingivalis haemaggl	8.72e+02
491	5	5.0	450	1	W69489	Haemagglutinin protein	8.72e+02
492	5	5.0	452	1	Y03766	S. aureus MurF polypep	8.72e+02
493	5	5.0	455	1	W37435	Rat hexokinase.	8.72e+02
494	5	5.0	456	1	R96022	P. gingivalis haemaggl	8.72e+02
495	5	5.0	456	1	R96023	P. gingivalis haemaggl	8.72e+02
496	5	5.0	456	1	W69490	Haemagglutinin protein	8.72e+02
497	5	5.0	456	1	R74996	E. maxima Em70-1 antig	8.72e+02
498	5	5.0	456	1	W69491	Haemagglutinin protein	8.72e+02
499	5	5.0	457	1	W98764	H. pylori GHPO 1103 pr	8.72e+02
500	5	5.0	458	1	W49908	Human brain derived ph	8.72e+02
501	5	5.0	461	1	W14007	Caenorhabditis elegans	8.72e+02
502	5	5.0	461	1	W82364	Flea saliva protein Pf	8.72e+02
503	5	5.0	461	1	W14006	Caenorhabditis elegans	8.72e+02
504	5	5.0	461	1	W33890	Flea saliva protein Pf	8.72e+02
505	5	5.0	464	1	W26496	CD2 associated intrace	8.72e+02
506	5	5.0	464	1	W80420	CD2 associated intrace	8.72e+02
507	5	5.0	464	1	W25116	CD2-associated intrace	8.72e+02
508	5	5.0	466	1	Y06981	Recombinant pHIV-210 p	8.72e+02
509	5	5.0	470	1	Y05398	Human TIE ligand NL8 p	8.72e+02
510	5	5.0	474	1	W00669	Glutathione synthetase	8.72e+02
511	5	5.0	474	1	R14676	Rabbit vitronectin-lik	8.72e+02
512	5	5.0	476	1	W46446	Human mature cholester	8.72e+02
513	5	5.0	476	1	R20540	Alkaline protease.	8.72e+02
514	5	5.0	476	1	W06127	Human cholesteryl este	8.72e+02
515	5	5.0	480	1	R12317	Chimeric HIV-2/HIV-1 e	8.72e+02
516	5	5.0	481	1	Y04145	Rat Tango-76 protein.	8.72e+02
517	5	5.0	488	1	R94579	Chlamydia pneumoniae p	8.72e+02
518	5	5.0	490	1	Y13366	Amino acid sequence of	8.72e+02
519	5	5.0	491	1	R60342	Partial human Lipid tr	8.72e+02
520	5	5.0	493	1	W99553	Protein encoded by clo	8.72e+02
521	5	5.0	493	1	R48669	Chitinase derivative #	8.72e+02
522	5	5.0	493	1	W71068	Multiple sclerosis ass	8.72e+02
523	5	5.0	493	1	W55989	Swinepox virus HindIII	8.72e+02
524	5	5.0	496	1	W76592	Simian immunodeficienc	8.72e+02
525	5	5.0	496	1	W76589	Simian immunodeficienc	8.72e+02
526	5	5.0	496	1	W76591	Simian immunodeficienc	8.72e+02
527	5	5.0	496	1	W76590	Simian immunodeficienc	8.72e+02
528	5	5.0	499	1	W17783	FIV integrase-LexA (F-	8.72e+02
529	5	5.0	500	1	W94496	Human delta-2 protein	8.72e+02
530	5	5.0	502	1	W99668	Human secreted protein	8.72e+02
531	5	5.0	504	1	Y00225	Enterococcus faecalis	8.72e+02
532	5	5.0	507	1	W21010	H. pylori chaperone pr	8.72e+02
533	5	5.0	507	1	W15574	Mouse Aiolos polypepti	8.72e+02
534	5	5.0	508	1	R30875	Prod. of the insert in	8.72e+02

535	5	5.0	511	1	W04539	Vesiculovirus glycopro	8.72e+02
536	5	5.0	511	1	W73506	VSV-G protein sequence	8.72e+02
537	5	5.0	512	1	W15274	Salmonella secreted pr	8.72e+02
538	5	5.0	513	1	R79945	Helicobacter pylori an	8.72e+02
539	5	5.0	517	1	R85852	WD-40 domain-contg. be	8.72e+02
540	5	5.0	519	1	R72589	A. crysogenum cystathi	8.72e+02
541	5	5.0	522	1	W78488	Human RIP140 ligand bi	8.72e+02
542	5	5.0	522	1	R66219	ADP-glucose-pyrophosph	8.72e+02
543	5	5.0	527	1	W47207	Homo sapiens tubulin-f	8.72e+02
544	5	5.0	527	1	W81560	Truncated RSV F protei	8.72e+02
545	5	5.0	527	1	W08367	RSV truncated F protei	8.72e+02
546	5	5.0	529	1	R97547	Arabidopsis adenylosuc	8.72e+02
547	5	5.0	530	1	W26428	Swinepox virus HindIII	8.72e+02
548	5	5.0	531	1	Y05376	Human HCMV inducible g	8.72e+02
549	5	5.0	532	1	W13127	Truncated rat cadherin	8.72e+02
550	5	5.0	532	1	W25640	Rat truncated cadherin	8.72e+02
551	5	5.0	535	1	W77299	Amino acid sequence of	8.72e+02
552	5	5.0	537	1	Y04661	L.lactis HsdM subunit	8.72e+02
553	5	5.0	538	1	R99854	Human OVCA1 tumour sup	8.72e+02
554	5	5.0	539	1	W25740	Cotton protoporphyrino	8.72e+02
555	5	5.0	539	1	W41608	Soybean protox-1.	8.72e+02
556	5	5.0	539	1	W20778	H. pylori cytoplasmic	8.72e+02
557	5	5.0	541	1	R54842	HER4 N-terminal trunca	8.72e+02
558	5	5.0	541	1	R91736	Receptor tyrosine kina	8.72e+02
559	5	5.0	541	1	W12695	G-protein parathyroid	8.72e+02
560	5	5.0	542	1	W58552	Human excitatory amino	8.72e+02
561	5	5.0	542	1	W26600	Human glutamate transp	8.72e+02
562	5	5.0	542	1	W83922	Human excitatory amino	8.72e+02
563	5	5.0	543	1	W20785	H. pylori cytoplasmic	8.72e+02
564	5	5.0	544	1	W28866	Rat brain Neuroglycan	8.72e+02
565	5	5.0	545	1	R41726	Peptide derived from H	8.72e+02
566	5	5.0	547	1	R67385	Mitochondrial protein	8.72e+02
567	5	5.0	550	1	Y13383	Amino acid sequence of	8.72e+02
568	5	5.0	551	1	R77872	SBP65 biotinylated pro	8.72e+02
569	5	5.0	551	1	R71376	Human cystathionine be	8.72e+02
570	5	5.0	552	1	Y00224	Enterococcus faecalis	8.72e+02
571	5	5.0	552	1	R26841	Cholesterol oxidase.	8.72e+02
572	5	5.0	553	1	W68489	Human partial ULIP-4 p	8.72e+02
573	5	5.0	553	1	W26495	CD2 associated intrace	8.72e+02
574	5	5.0	553	1	W25115	CD2-associated intrace	8.72e+02
575	5	5.0	553	1	W80419	CD2 associated intrace	8.72e+02
576	5	5.0	558	1	Y03633	Hypoxia-regulated gene	8.72e+02
577	5	5.0	558	1	W14502	Saccharomyces cerevisi	8.72e+02
578	5	5.0	558	1	P70429	Murine neuroleukin.	8.72e+02
579	5	5.0	559	1	W89803	Staphylococcus aureus	8.72e+02
580	5	5.0	560	1	W70500	Human sodium-lithium c	8.72e+02
581	5	5.0	561	1	W97746	Human endogenous retro	8.72e+02
582	5	5.0	561	1	W95693	Human endogenous retro	8.72e+02
583	5	5.0	562	1	R63588	Full length HA protein	8.72e+02
584	5	5.0	566	1	R08259	Haemagglutinin.	8.72e+02
585	5	5.0	566	1	W04271	B.t. neutral protease.	8.72e+02
586	5	5.0	566	1	W13733	Protease NprL.	8.72e+02
587	5	5.0	566	1	W68405	SIV strain H1N1 haemag	8.72e+02
588	5	5.0	567	1	R25310	Bovine RSV strain FS-i	8.72e+02
589	5	5.0	568	1	W03325	Newcastle disease viru	8.72e+02
590	5	5.0	568	1	W49683	Newcastle disease viru	8.72e+02
591	5	5.0	568	1	W03551	Newcastle disease viru	8.72e+02
592	5	5.0	568	1	W62989	HN protein of Newcastl	8.72e+02
593	5	5.0	571	1	W55997	Protein SEQ ID NO:227	8.72e+02
594	5	5.0	571	1	P91960	Haemagglutinin and neu	8.72e+02
595	5	5.0	571	1	W44940	Newcastle disease viru	8.72e+02
596	5	5.0	572	1	R24189	Bovine RSV strain A 51	8.72e+02
597	5	5.0	572	1	W75442	Influenza virus A/Texa	8.72e+02
598	5	5.0	572	1	W68488	Mouse ULIP-4 protein.	8.72e+02
599	5	5.0	572	1	W01670	Influenza A/Texas/36/9	8.72e+02
600	5	5.0	573	1	W00164	Myosin heavy chain ana	8.72e+02
601	5	5.0	574	1	W47604	HRSV glycoprotein F.	8.72e+02
602	5	5.0	574	1	P93562	Sequence of respirator	8.72e+02
603	5	5.0	574	1	W81559	Respiratory syncytial	8.72e+02
604	5	5.0	574	1	R21327	Sequence of protein F.	8.72e+02
605	5	5.0	574	1	R79894	RS virus fusion protei	8.72e+02
606	5	5.0	574	1	R39285	Respiratory syncitial	8.72e+02
607	5	5.0	574	1	R25301	HRSV glycoprotein F (g	8.72e+02
608	5	5.0	575	1	Y00157	Enterococcus faecalis	8.72e+02
609	5	5.0	577	1	W06827	Newcastle disease viru	8.72e+02
610	5	5.0	577	1	R39703	Haemagglutinin-neurami	8.72e+02
611	5	5.0	577	1	R80558	Newcastle's disease vi	8.72e+02
612	5	5.0	578	1	Y09065	Human complement facto	8.72e+02
613	5	5.0	579	1	R34401	Sequence of low molecu	8.72e+02
614	5	5.0	579	1	R34446	N.meningitidis 2394 Tb	8.72e+02
615	5	5.0	580	1	W02164	Lactococcus lactis sub	8.72e+02
616	5	5.0	581	1	R58598	Newcastle disease viru	8.72e+02
617	5	5.0	581	1	W10690	Newcastle disease viru	8.72e+02
618	5	5.0	581	1	R58859	Newcastle Disease Viru	8.72e+02
619	5	5.0	587	1	Y11082	H. pylori ORF hp4p6285	8.72e+02
620	5	5.0	587	1	Y03204	Amino acid sequence of	8.72e+02
621	5	5.0	590	1	Y11093	H. pylori ORF hp6p1224	8.72e+02
622	5	5.0	592	1	R26420	Cyclodextrinase.	8.72e+02
623	5	5.0	593	1	W71506	Helicobacter polypepti	8.72e+02
624	5	5.0	594	1	W48089	Aeromonas caviae polye	8.72e+02
625	5	5.0	598	1	R12319	HIV-2/HIV-1/HTLV-I chi	8.72e+02
626	5	5.0	599	1	Y01528	TbpB sequence from Nei	8.72e+02
627	5	5.0	599	1	R88644	Neisseria meningitidis	8.72e+02
628	5	5.0	599	1	R48219	N.meningitidis IM2394	8.72e+02
629	5	5.0	599	1	W14620	HTR Tbp2 protein from	8.72e+02
630	5	5.0	600	1	R43003	Mycoplasma 74.5kD prot	8.72e+02
631	5	5.0	600	1	R09419	Hsp70 antigen of Mycop	8.72e+02
632	5	5.0	600	1	R03922	M.hyopneumoniae HSP (M	8.72e+02
633	5	5.0	601	1	Y00156	Enterococcus faecalis	8.72e+02
634	5	5.0	603	1	Y05277	P. roqueforti beta-fru	8.72e+02
635	5	5.0	609	1	W09421	Lettuce polyphenol oxi	8.72e+02
636	5	5.0	616	1	P96147	Sequence of Newcastle	8.72e+02
637	5	5.0	616	1	R06329	Newcastle disease viru	8.72e+02
638	5	5.0	621	1	W78420	Arabidopsis thaliana e	8.72e+02
639	5	5.0	621	1	R41230	Taurine transporter.	8.72e+02
640	5	5.0	622	1	R67424	Alginic acid lyase.	8.72e+02
641	5	5.0	626	1	W37140	Pea plastidial phospho	8.72e+02
642	5	5.0	626	1	R71034	N. crassa glucoamylase	8.72e+02
643	5	5.0	627	1	Y00235	Enterococcus faecalis	8.72e+02
644	5	5.0	633	1	W34470	Human N-acetylgalactos	8.72e+02
645	5	5.0	635	1	W55723	H. pylori ORF hp4e5339	8.72e+02
646	5	5.0	637	1	W77034	Artichoke sucrose fruc	8.72e+02
647	5	5.0	637	1	R63673	Glutamylcysteine-synth	8.72e+02
648	5	5.0	637	1	P92108	Human growth hormone r	8.72e+02
649	5	5.0	638	1	P92107	Rabbit growth hormone	8.72e+02
650	5	5.0	638	1	W33395	Rabbit growth hormone	8.72e+02
651	5	5.0	638	1	W33394	Human growth hormone r	8.72e+02
652	5	5.0	638	1	P81327	Rabbit growth hormone	8.72e+02
653	5	5.0	638	1	W86014	R. rubrum poly-beta-hy	8.72e+02
654	5	5.0	638	1	P81326	Human growth hormone r	8.72e+02
655	5	5.0	639	1	R47172	Sequence of the inner	8.72e+02
656	5	5.0	646	1	W55314	H. pylori ORF hp1p1385	8.72e+02
657	5	5.0	647	1	W73376	Human HPDDV78 protein	8.72e+02
658	5	5.0	649	1	R94584	DHFR/C. pneumoniae ant	8.72e+02
659	5	5.0	652	1	R88124	Tobacco mosaic virus r	8.72e+02
660	5	5.0	657	1	W32095	Porcine retrovirus par	8.72e+02
661	5	5.0	659	1	W94497	Human delta-2 mature p	8.72e+02
662	5	5.0	660	1	W98211	H. pylori GHPO 761 pro	8.72e+02
663	5	5.0	660	1	W71492	Helicobacter polypepti	8.72e+02
664	5	5.0	660	1	W85453	Pig endogenous retrovi	8.72e+02
665	5	5.0	661	1	R07504	Merozite apical-end-lo	8.72e+02
666	5	5.0	662	1	W24576	Merozoite apical-end p	8.72e+02
667	5	5.0	665	1	R36805	Pseudomonas exotoxin-i	8.72e+02
668	5	5.0	667	1	W20753	H. pylori transporter	8.72e+02
669	5	5.0	668	1	P81186	Sequence encoded by ne	8.72e+02
670	5	5.0	668	1	W55709	H. pylori ORF hp4e5339	8.72e+02
671	5	5.0	668	1	P82924	Cowpox virus protein i	8.72e+02
672	5	5.0	672	1	R66726	Protein kinase C mutan	8.72e+02
673	5	5.0	672	1	R94765	Type III (alpha-type)	8.72e+02
674	5	5.0	678	1	R99414	Human gas6 protein, an	8.72e+02
675	5	5.0	678	1	W98387	H. pylori GHPO 16 prot	8.72e+02
676	5	5.0	678	1	W46463	Human growth arrest sp	8.72e+02
677	5	5.0	681	1	P90441	Chimeric human respira	8.72e+02
678	5	5.0	683	1	R80573	Human beta-IG-H3 (tran	8.72e+02
679	5	5.0	683	1	R40386	betaIG-H3 protein.	8.72e+02
680	5	5.0	685	1	W94507	Human delta-2 protein.	8.72e+02

681	5	5.0	685	1	W80813	Nucleotide sequence of	8.72e+02
682	5	5.0	687	1	W22458	Yeast wild-type acetol	8.72e+02
683	5	5.0	687	1	P81152	Herbicide sensitive, w	8.72e+02
684	5	5.0	688	1	W43107	C. thermocellum ORF2p	8.72e+02
685	5	5.0	688	1	W06316	TRP-1 protein.	8.72e+02
686	5	5.0	702	1	W72184	HSV-2 strain SB5 Conti	8.72e+02
687	5	5.0	705	1	W79144	Receptor protein tyros	8.72e+02
688	5	5.0	705	1	W81401	Receptor protein tyros	8.72e+02
689	5	5.0	707	1	W41523	Blood group antigen bi	8.72e+02
690	5	5.0	708	1	R71037	Draculin.	8.72e+02
691	5	5.0	708	1	W73022	Helicobacter pylori 76	8.72e+02
692	5	5.0	709	1	W58988	Homo sapiens fetal kid	8.72e+02
693	5	5.0	719	1	R45775	Murine leukaemia inhib	8.72e+02
694	5	5.0	719	1	R25069	mLIF-R.	8.72e+02
695	5	5.0	719	1	R49507	Human LIF-R clone 65.	8.72e+02
696	5	5.0	719	1	R74096	Murine leukaemia inhib	8.72e+02
697	5	5.0	720	1	W19266	Lactobacillus amylovor	8.72e+02
698	5	5.0	722	1	W55104	Streptococcus pneumoni	8.72e+02
699	5	5.0	723	1	W94905	Mouse pheromone recept	8.72e+02
700	5	5.0	739	1	W31549	Bloom's syndrome BLM m	8.72e+02
[redacted]	5	5.0	741	1	W22519	Glycosyl transferase.	8.72e+02
[redacted]	5	5.0	744	1	W19907	Transmembrane glycopro	8.72e+02
703	5	5.0	745	1	P90615	Polypeptide with isoam	8.72e+02
704	5	5.0	747	1	W60487	Human TRIDENT transcri	8.72e+02
705	5	5.0	748	1	W55725	H. pylori ORF 05ae3022	8.72e+02
706	5	5.0	748	1	R24513	Consensus sequence enc	8.72e+02
707	5	5.0	748	1	R24398	Prod. of the S gene of	8.72e+02
708	5	5.0	748	1	R24396	Prod. of the S gene of	8.72e+02
709	5	5.0	748	1	W55511	H. pylori ORF 05ae3022	8.72e+02
710	5	5.0	748	1	W98877	H. pylori GHPO 1750 pr	8.72e+02
711	5	5.0	749	1	R42466	Feline infectious peri	8.72e+02
712	5	5.0	750	1	P94419	Primary transcript of	8.72e+02
713	5	5.0	751	1	W61119	Neutral trehalase 1 (N	8.72e+02
714	5	5.0	755	1	W03550	Partial ORF-3 protein	8.72e+02
715	5	5.0	756	1	R08405	AmF105L encoded by Flo	8.72e+02
716	5	5.0	757	1	W60486	Mouse TRIDENT transcri	8.72e+02
717	5	5.0	758	1	R86998	Enterohaemorrhagic E.c	8.72e+02
718	5	5.0	763	1	W55095	Streptococcus pneumoni	8.72e+02
719	5	5.0	763	1	W60591	Human hepatocyte nucle	8.72e+02
720	5	5.0	766	1	W13385	Human protein ubiquiti	8.72e+02
721	5	5.0	768	1	W36823	Human E2A-binding prot	8.72e+02
722	5	5.0	768	1	W36815	Human E2A-binding prot	8.72e+02
723	5	5.0	771	1	W60590	Rat hepatocyte nuclear	8.72e+02
724	5	5.0	776	1	P71568	Sequence of rotavirus	8.72e+02
725	5	5.0	776	1	R24293	VP4 protein of strain	8.72e+02
726	5	5.0	789	1	R15143	Haemonchus contortus p	8.72e+02
727	5	5.0	793	1	W25635	Human cadherin-8.	8.72e+02
728	5	5.0	793	1	W13132	Full length human cadh	8.72e+02
[redacted]	5	5.0	796	1	W55090	Streptococcus pneumoni	8.72e+02
[redacted]	5	5.0	799	1	W25639	Rat full length cadher	8.72e+02
731	5	5.0	799	1	W13126	Full length rat cadher	8.72e+02
732	5	5.0	804	1	W94914	Mouse pheromone recept	8.72e+02
733	5	5.0	810	1	W93003	Murine IL-4 receptor c	8.72e+02
734	5	5.0	810	1	W60669	Mouse interleukin-4 re	8.72e+02
735	5	5.0	810	1	W73471	Mouse interleukin-4 re	8.72e+02
736	5	5.0	810	1	W13498	Murine interleukin-4 r	8.72e+02
737	5	5.0	810	1	R11255	Murine IL-4 receptor.	8.72e+02
738	5	5.0	810	1	W48313	Mouse interleukin-4 re	8.72e+02
739	5	5.0	810	1	R04574	Derived amino acid seq	8.72e+02
740	5	5.0	821	1	R35451	Mouse eps8.	8.72e+02
741	5	5.0	829	1	R27549	MCC gene product.	8.72e+02
742	5	5.0	841	1	W88420	Chlamydia pneumoniae s	8.72e+02
743	5	5.0	842	1	R93804	Class A starch branchi	8.72e+02
744	5	5.0	843	1	R67760	Lys-aminopeptidase Pep	8.72e+02
745	5	5.0	845	1	W36816	Human E2A-binding prot	8.72e+02
746	5	5.0	847	1	W06400	Class A starch branchi	8.72e+02
747	5	5.0	850	1	W57445	A. thermophilum thermo	8.72e+02
748	5	5.0	852	1	W94904	Mouse pheromone recept	8.72e+02
749	5	5.0	852	1	R04029	Env gene product of cD	8.72e+02
750	5	5.0	856	1	P80803	Sequence of env protei	8.72e+02
751	5	5.0	857	1	W13052	HIV-2 provirus-encoded	8.72e+02
752	5	5.0	858	1	P80806	Sequence of env protei	8.72e+02
753	5	5.0	858	1	P81779	Sequence encoded by op	8.72e+02
754	5	5.0	861	1	W69977	Staphylococcus aureus	8.72e+02
755	5	5.0	864	1	R24042	Lipoxygenase.	8.72e+02
756	5	5.0	878	1	Y00868	S. tuberosum isoamylas	8.72e+02
757	5	5.0	878	1	W19113	Potato starch branchin	8.72e+02
758	5	5.0	879	1	W89320	SIVmac239 genome env p	8.72e+02
759	5	5.0	879	1	R24126	SIVmac239 env gene pro	8.72e+02
760	5	5.0	882	1	R43959	N-heparan sulphate sul	8.72e+02
761	5	5.0	882	1	W06399	Class A starch branchi	8.72e+02
762	5	5.0	891	1	P82677	ENVRN sequence from HI	8.72e+02
763	5	5.0	906	1	R44300	Human GluR1B receptor.	8.72e+02
764	5	5.0	906	1	R84913	Human glutamate recept	8.72e+02
765	5	5.0	906	1	R42354	Human glutamate recept	8.72e+02
766	5	5.0	906	1	R84912	Human glutamate recept	8.72e+02
767	5	5.0	906	1	W36795	Novel human protein, d	8.72e+02
768	5	5.0	907	1	R11989	Glutamate receptor 1.	8.72e+02
769	5	5.0	910	1	W69248	Clone CT797_3 protein	8.72e+02
770	5	5.0	916	1	W85023	Amino acid sequence of	8.72e+02
771	5	5.0	917	1	R64974	Pol region gene produc	8.72e+02
772	5	5.0	918	1	W37428	Rat hexokinase I.	8.72e+02
773	5	5.0	918	1	W37442	Rat hexokinase I.	8.72e+02
774	5	5.0	919	1	R50179	Excitatory amino acid	8.72e+02
775	5	5.0	919	1	W37440	Hexokinase-glucokinase	8.72e+02
776	5	5.0	919	1	W37433	Hexokinase-glucokinase	8.72e+02
777	5	5.0	919	1	W37436	Hexokinase-glucokinase	8.72e+02
778	5	5.0	919	1	W18580	Potato alpha-glucosida	8.72e+02
779	5	5.0	921	1	W60752	cGS-PDE isolated from	8.72e+02
780	5	5.0	921	1	W77040	Adrenal cortex Ca2+/ca	8.72e+02
781	5	5.0	921	1	W18048	Cyclic-GMP-stimulated	8.72e+02
782	5	5.0	921	1	W71224	cGS-PDE encoded by p3C	8.72e+02
783	5	5.0	921	1	R69727	Cyclic-GMP stimulated	8.72e+02
784	5	5.0	921	1	R28407	cGS-PDE from clone 3CG	8.72e+02
785	5	5.0	923	1	W96246	Human semaphorin recep	8.72e+02
786	5	5.0	923	1	W96248	Mouse semaphorin recep	8.72e+02
787	5	5.0	928	1	W88421	Chlamydia pneumoniae s	8.72e+02
788	5	5.0	928	1	W88423	Chlamydia pneumoniae s	8.72e+02
789	5	5.0	930	1	W88424	Chlamydia pneumoniae s	8.72e+02
790	5	5.0	940	1	R07070	Fibronectin-binding pr	8.72e+02
791	5	5.0	941	1	W18050	Human cGS-PDE amino ac	8.72e+02
792	5	5.0	941	1	W11253	pHcgs6n cyclic GMP sti	8.72e+02
793	5	5.0	941	1	W77042	Cyclic-GMP-stimulated	8.72e+02
794	5	5.0	941	1	W60754	Human cyclic-GMP-nucle	8.72e+02
795	5	5.0	941	1	W71226	Human cGS-PDE cDNA pro	8.72e+02
796	5	5.0	941	1	R28409	Human foetal cGS PDE f	8.72e+02
797	5	5.0	942	1	R28408	CGS PDE from bovine br	8.72e+02
798	5	5.0	942	1	W11239	Cyclic GMP stimulated	8.72e+02
799	5	5.0	942	1	W18049	cGS-PDE amino acid seq	8.72e+02
800	5	5.0	942	1	W60753	cGS-PDE isolated from	8.72e+02
801	5	5.0	942	1	W77041	Cyclic-GMP-stimulated	8.72e+02
802	5	5.0	942	1	W71225	Bovine brain cGS-PDE s	8.72e+02
803	5	5.0	942	1	R69728	Cyclic-GMP stimulated	8.72e+02
804	5	5.0	946	1	W27085	Human cytomegalovirus	8.72e+02
805	5	5.0	946	1	W27084	Human cytomegalovirus	8.72e+02
806	5	5.0	950	1	R06376	Product of the ssc1 ge	8.72e+02
807	5	5.0	954	1	W19752	Yeast inhibitor of apo	8.72e+02
808	5	5.0	956	1	W49874	Bankia gouldi glycosid	8.72e+02
809	5	5.0	958	1	W19969	Yeast alanyl-tRNA synt	8.72e+02
810	5	5.0	963	1	W70540	Integrin beta-1 chain.	8.72e+02
811	5	5.0	973	1	W41815	A. thaliana trehalose	8.72e+02
812	5	5.0	975	1	W72746	Drosophila kinesin.	8.72e+02
813	5	5.0	977	1	W59881	Amino acid sequence of	8.72e+02
814	5	5.0	985	1	R42995	Glycosyltransferase.	8.72e+02
815	5	5.0	985	1	R42214	Aspergillus niger gluc	8.72e+02
816	5	5.0	986	1	W59360	C.elegans retinal dege	8.72e+02
817	5	5.0	990	1	W37786	Amino acid sequence of	8.72e+02
818	5	5.0	990	1	R84660	Chondroitinase II.	8.72e+02
819	5	5.0	990	1	W90075	P. vulgaris chondroiti	8.72e+02
820	5	5.0	990	1	W39746	P. vulgaris chondroiti	8.72e+02
821	5	5.0	990	1	R77460	Chondroitinase-II.	8.72e+02
822	5	5.0	993	1	R95268	Pre-nisin modification	8.72e+02
823	5	5.0	995	1	R84658	STLVpan-p pol protein.	8.72e+02
824	5	5.0	999	1	W98498	H. pylori GHPO 1115 pr	8.72e+02
825	5	5.0	1013	1	W09398	Proteus vulgaris chond	8.72e+02
826	5	5.0	1013	1	R62535	P. vulgaris chondroiti	8.72e+02

827	5	5.0	1013	1	W40224	Human tolloid-like (Tl	8.72e+02
828	5	5.0	1013	1	W40223	Murine mTll protein.	8.72e+02
829	5	5.0	1015	1	W63675	Polypeptide having age	8.72e+02
830	5	5.0	1028	1	W29667	Homo sapiens DL185_1 c	8.72e+02
831	5	5.0	1031	1	W79062	FIV-NCSU1 clone JSY3 p	8.72e+02
832	5	5.0	1038	1	W36903	Human epididymis-speci	8.72e+02
833	5	5.0	1058	1	R91734	Receptor tyrosine kina	8.72e+02
834	5	5.0	1058	1	R54843	HER4 with alternate 3'	8.72e+02
835	5	5.0	1088	1	R88635	Plasmid pASK75 open re	8.72e+02
836	5	5.0	1093	1	R66460	AF-17 protein.	8.72e+02
837	5	5.0	1095	1	P97052	Sequence encoded by in	8.72e+02
838	5	5.0	1095	1	W80359	An F-actin-combined pr	8.72e+02
839	5	5.0	1124	1	W53668	FIV PPR clone 34 ORF2	8.72e+02
840	5	5.0	1124	1	R24237	Pol polypeptide of FIV	8.72e+02
841	5	5.0	1128	1	W36817	Mouse E2A-binding prot	8.72e+02
842	5	5.0	1128	1	R49994	Mouse carboxypeptidase	8.72e+02
843	5	5.0	1136	1	P93341	Sequence encoded by to	8.72e+02
844	5	5.0	1136	1	P82314	Bacillus thuringiensis	8.72e+02
845	5	5.0	1137	1	R42081	Impatiens Necrotic Spo	8.72e+02
846	5	5.0	1140	1	R72386	XAP-1, part of the DNA	8.72e+02
847	5	5.0	1144	1	R76059	Mycoplasma pirum adhes	8.72e+02
848	5	5.0	1144	1	R88123	Tobacco mosaic virus r	8.72e+02
849	5	5.0	1144	1	R88122	Tobacco mosaic virus r	8.72e+02
850	5	5.0	1151	1	W95039	Human N-arginine dibas	8.72e+02
851	5	5.0	1155	1	W60002	Mouse alpha d polypeti	8.72e+02
852	5	5.0	1155	1	W65102	Mouse beta-integrin al	8.72e+02
853	5	5.0	1155	1	W73346	Mouse alphad protein s	8.72e+02
854	5	5.0	1155	1	W72835	Mouse alpha-d #1.	8.72e+02
855	5	5.0	1155	1	W23060	Mouse beta 2 integrin	8.72e+02
856	5	5.0	1155	1	R78167	Mouse alpha-d subunit.	8.72e+02
857	5	5.0	1161	1	W23061	Mouse beta 2 integrin	8.72e+02
858	5	5.0	1161	1	W65103	Mouse beta-integrin al	8.72e+02
859	5	5.0	1161	1	W73347	Mouse alphad protein s	8.72e+02
860	5	5.0	1161	1	W72836	Mouse alpha-d #2.	8.72e+02
861	5	5.0	1161	1	R78168	Mouse alpha-d subunit.	8.72e+02
862	5	5.0	1161	1	W60003	Mouse alpha d polypeti	8.72e+02
863	5	5.0	1167	1	W31504	Nematode toxin 167P pr	8.72e+02
864	5	5.0	1167	1	W10653	Bacillus thuringiensis	8.72e+02
865	5	5.0	1168	1	W16326	Nematocidal toxin 167P	8.72e+02
866	5	5.0	1178	1	P60051	Sequence of insecticid	8.72e+02
867	5	5.0	1186	1	W21694	Bacillus thuringiensis	8.72e+02
868	5	5.0	1186	1	W31503	Nematode toxin 80JJ1 p	8.72e+02
869	5	5.0	1186	1	R56449	Delta endotoxin from B	8.72e+02
870	5	5.0	1186	1	W10652	Bacillus thuringiensis	8.72e+02
871	5	5.0	1194	1	W91071	Apoptosis inducer Apaf	8.72e+02
872	5	5.0	1203	1	W81572	Mus dunni endogenous v	8.72e+02
873	5	5.0	1203	1	W57315	Human endothelial citr	8.72e+02
874	5	5.0	1203	1	R41668	Human endothelial cell	8.72e+02
875	5	5.0	1205	1	W91072	Apoptosis inducer spli	8.72e+02
876	5	5.0	1205	1	R77363	Endothelial nitrogen m	8.72e+02
877	5	5.0	1221	1	R52699	Sequence translated fr	8.72e+02
878	5	5.0	1253	1	W88524	Adenyl cyclase-1 (acy-	8.72e+02
879	5	5.0	1277	1	R52700	Plasmid pASK60-Strep r	8.72e+02
880	5	5.0	1287	1	R79944	Helicobacter pylori va	8.72e+02
881	5	5.0	1288	1	W55685	H. pylori ORF 07ee1140	8.72e+02
882	5	5.0	1288	1	W55547	H. pylori ORF 14ee4192	8.72e+02
883	5	5.0	1290	1	W98269	H. pylori GHPO 374 pro	8.72e+02
884	5	5.0	1294	1	W30601	Human type IX adenylyl	8.72e+02
885	5	5.0	1296	1	R41198	CT.	8.72e+02
886	5	5.0	1308	1	R91733	Receptor tyrosine kina	8.72e+02
887	5	5.0	1308	1	R54841	HER4.	8.72e+02
888	5	5.0	1311	1	W72971	Precis coenia patched	8.72e+02
889	5	5.0	1311	1	W52197	Precis coenia (butterf	8.72e+02
890	5	5.0	1313	1	Y00234	Enterococcus faecalis	8.72e+02
891	5	5.0	1313	1	Y00236	Enterococcus faecalis	8.72e+02
892	5	5.0	1358	1	R99358	Human restrictin.	8.72e+02
893	5	5.0	1358	1	W18824	Human restrictin.	8.72e+02
894	5	5.0	1358	1	W35743	Recombinant human rest	8.72e+02
895	5	5.0	1375	1	W27283	Apoptosis inducing pro	8.72e+02
896	5	5.0	1382	1	Y11001	H. pylori ORF 11ee1140	8.72e+02
897	5	5.0	1385	1	R28889	Toxin 17a.	8.72e+02
898	5	5.0	1385	1	R44201	Bacillus thuringiensis	8.72e+02
899	5	5.0	1385	1	W13884	17a toxin.	8.72e+02
900	5	5.0	1385	1	R29516	BT toxin 17a.	8.72e+02
901	5	5.0	1385	1	R28803	BT toxin 17a.	8.72e+02
902	5	5.0	1385	1	R29026	Bacillus thuringiensis	8.72e+02
903	5	5.0	1385	1	R20066	B.thuringiensis toxin	8.72e+02
904	5	5.0	1385	1	R76112	PS17a acaride-active t	8.72e+02
905	5	5.0	1385	1	R58631	Bacillus thuringiensis	8.72e+02
906	5	5.0	1392	1	Y06999	Restin protein sequenc	8.72e+02
907	5	5.0	1400	1	R92705	MLL cDNA clone 14-18B	8.72e+02
908	5	5.0	1400	1	R44514	MLL amino acid sequenc	8.72e+02
909	5	5.0	1413	1	W20725	H. pylori secreted or	8.72e+02
910	5	5.0	1416	1	Y00211	Enterococcus faecalis	8.72e+02
911	5	5.0	1417	1	W31551	Bloom's syndrome BLM m	8.72e+02
912	5	5.0	1417	1	W31548	Bloom's syndrome BLM m	8.72e+02
913	5	5.0	1417	1	W31550	Bloom's syndrome BLM m	8.72e+02
914	5	5.0	1423	1	R74205	Human death associated	8.72e+02
915	5	5.0	1423	1	W71367	Death associated prote	8.72e+02
916	5	5.0	1427	1	R10534	Human 160kD mediator o	8.72e+02
917	5	5.0	1437	1	W80597	Human multidrug resist	8.72e+02
918	5	5.0	1448	1	Y00210	Enterococcus faecalis	8.72e+02
919	5	5.0	1454	1	R42473	FIPV/FECV chimeric spi	8.72e+02
920	5	5.0	1454	1	R42478	FECV/FIPV chimeric spi	8.72e+02
921	5	5.0	1454	1	R42464	Feline infectious peri	8.72e+02
922	5	5.0	1454	1	R24395	Prod. of the S gene of	8.72e+02
923	5	5.0	1454	1	R24511	Prod. of the S genome	8.72e+02
924	5	5.0	1454	1	R42475	FECV/FIPV chimeric spi	8.72e+02
925	5	5.0	1454	1	R42467	Feline infectious peri	8.72e+02
926	5	5.0	1454	1	R42468	Feline enteric coronav	8.72e+02
927	5	5.0	1454	1	R42477	FECV/FIPV chimeric spi	8.72e+02
928	5	5.0	1454	1	R42470	Feline infectious peri	8.72e+02
929	5	5.0	1454	1	R42472	FECV/FIPV chimeric spi	8.72e+02
930	5	5.0	1454	1	R24397	Prod. of the S gene of	8.72e+02
931	5	5.0	1454	1	R42474	FECV/FIPV chimeric spi	8.72e+02
932	5	5.0	1456	1	R49042	NMDA receptor channel	8.72e+02
933	5	5.0	1461	1	W64468	Human secreted protein	8.72e+02
934	5	5.0	1464	1	W64229	Potato starch-associat	8.72e+02
935	5	5.0	1464	1	W14920	Berolina potato starch	8.72e+02
936	5	5.0	1464	1	W18792	Starch binding protein	8.72e+02
937	5	5.0	1482	1	R44193	Rat NMDA receptor subu	8.72e+02
938	5	5.0	1484	1	W55686	H. pylori ORF 07ee5070	8.72e+02
939	5	5.0	1484	1	W89721	Canine ribosome recept	8.72e+02
940	5	5.0	1594	1	P81183	Sequence of the peplom	8.72e+02
941	5	5.0	1614	1	W83312	Mouse Lrp5 protein.	8.72e+02
942	5	5.0	1644	1	W25049	BRCA2 cancer susceptib	8.72e+02
943	5	5.0	1671	1	R06341	Signal peptide and fir	8.72e+02
944	5	5.0	1732	1	W24787	PrtK antigenic protein	8.72e+02
945	5	5.0	1732	1	W69487	Haemagglutinin protein	8.72e+02
946	5	5.0	1732	1	R96029	P. gingivalis porphypa	8.72e+02
947	5	5.0	1757	1	W84351	Murine ubiquitin-prote	8.72e+02
948	5	5.0	1849	1	W56573	Toxin TcdAii, encoded	8.72e+02
949	5	5.0	1946	1	W47273	Lactobacillus bulgaric	8.72e+02
950	5	5.0	1969	1	W72419	Rice bacterial leaf bl	8.72e+02
951	5	5.0	2008	1	W22016	Utrophin truncated pol	8.72e+02
952	5	5.0	2016	1	W23994	Human hH1 sodium chann	8.72e+02
953	5	5.0	2019	1	R67913	Cardiac sodium channel	8.72e+02
954	5	5.0	2020	1	R06584	Cardiac sodium channel	8.72e+02
955	5	5.0	2179	1	P60243	Sequence encoding the	8.72e+02
956	5	5.0	2189	1	R05222	Antigen GX5401FL encod	8.72e+02
957	5	5.0	2329	1	W25038	Partial BRCA2 cancer s	8.72e+02
958	5	5.0	2368	1	W26663	Yeast checkpoint contr	8.72e+02
959	5	5.0	2368	1	W73893	Yeast MEC1 protein seq	8.72e+02
960	5	5.0	2408	1	R24306	Translation of ORF 2 c	8.72e+02
961	5	5.0	2471	1	Y06816	Human Notch2 (humN2) p	8.72e+02
962	5	5.0	2516	1	W17899	Photorhabdus luminesce	8.72e+02
963	5	5.0	2516	1	W56572	Toxin TcdA, encoded by	8.72e+02
964	5	5.0	2625	1	W55887	Human telomerase.	8.72e+02
965	5	5.0	2627	1	W61347	Human telomerase RNA i	8.72e+02
966	5	5.0	2628	1	R96030	P. gingivalis haemaggl	8.72e+02
967	5	5.0	2628	1	W69488	Haemagglutinin protein	8.72e+02
968	5	5.0	2629	1	W55885	Rat telomerase.	8.72e+02
969	5	5.0	2629	1	W61348	Mouse telomerase RNA i	8.72e+02
970	5	5.0	2721	1	P70647	Sequence of N-terminal	8.72e+02
971	5	5.0	2861	1	W27227	Human TRIO phosphoprot	8.72e+02
972	5	5.0	2893	1	W71556	Helicobacter polypepti	8.72e+02

973	5	5.0	2893	1	W98828	H. pylori GHPO 1484 pr	8.72e+02
974	5	5.0	2938	1	R59923	GAP protein Iral.	8.72e+02
975	5	5.0	2963	1	W56444	Fragment HGJ1789 of a	8.72e+02
976	5	5.0	3014	1	R35207	Hepatitis C virus prot	8.72e+02
977	5	5.0	3014	1	R54099	NANBHV E1/E2 protein.	8.72e+02
978	5	5.0	3066	1	W36178	Murine Ataxia-telangie	8.72e+02
979	5	5.0	3077	1	P93283	Sequence of clone HIV-	8.72e+02
980	5	5.0	3211	1	P81769	Sequence encoded by to	8.72e+02
981	5	5.0	3224	1	W54235	Human Nup358 protein.	8.72e+02
982	5	5.0	3418	1	Y04357	Human BRCA2 (omi4) pro	8.72e+02
983	5	5.0	3418	1	Y04356	Human BRCA2 (omi3) pro	8.72e+02
984	5	5.0	3418	1	Y04358	Human BRCA2 (omi5) pro	8.72e+02
985	5	5.0	3418	1	Y04355	Human BRCA2 (omi2) pro	8.72e+02
986	5	5.0	3418	1	Y04354	Human BRCA2 (omi1) pro	8.72e+02
987	5	5.0	3418	1	W23287	Human breast and ovari	8.72e+02
988	5	5.0	3418	1	W19211	Human breast cancer su	8.72e+02
989	5	5.0	3433	1	W22017	Utrophin.	8.72e+02
990	5	5.0	3588	1	R34712	Bacillus subtilis srfA	8.72e+02
991	5	5.0	3910	1	R66462	ALL-1 (acute lymphocyt	8.72e+02
992	5	5.0	3969	1	R52971	Product of the cDNA en	8.72e+02
[illegible]	5	5.0	4472	1	W22601	Tylactone synthase ORF	8.72e+02
[illegible]	5	5.0	4536	1	W96826	Amino acid sequence of	8.72e+02
995	5	5.0	4536	1	W41262	Apolipoprotein B-100.	8.72e+02
996	5	5.0	4545	1	W22611	Hybrid srmG/tylG ORF1	8.72e+02
997	5	5.0	4550	1	W22606	Platenolide synthase O	8.72e+02
998	5	5.0	4550	1	W23716	Platenolide synthase O	8.72e+02
999	5	5.0	4866	1	W77410	Human ryanodin recepto	8.72e+02
1000	5	5.0	15281	1	R44929	T. niveum Cyclosporin	8.72e+02

ALIGNMENTS

```
RESULT    1
ID   W80552 standard; Protein; 101 AA.
AC   W80552;
DT   09-DEC-1998  (first entry)
DE   Antiviral protein, cyanovirin-N.
KW   Cyanovirin-N; recombinant; cyanovirin; cyanobacterium; antiviral;
KW   HIV-1.
OS   Nostoc ellipsosporum.
PN   US5821081-A.
PD   13-OCT-1998.
PF   26-APR-1996; 638610.
PR   26-APR-1996; US-638610.
PR   27-APR-1995; US-429965.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI   Boyd MR, Gustafson KR, McMahon JB, Shoemaker RH;
DR   WPI; 98-567657/48.
DR   N-PSDB; V56025.
PT   DNA encoding cyanovirin polypeptide(s) - useful for producing
PT   recombinant polypeptides with antiviral activity
PS   Claim 1; Fig 2; 33pp; English.
CC   This represents an antiviral protein cyanovirin-N. A vector containing
CC   a nucleic acid molecule (V56025 or V56026) can be used to transform a
CC   host cell for the recombinant production of the cyanovirin polypeptides.
CC   The cyanovirins are derived from the cyanobacterium Nostoc ellipsosporum,
CC   and have antiviral activity, e.g. with EC50 values of 0.4-7.6 nM against
CC   various HIV-1 strains and isolates.
SQ   Sequence   101 AA;

  Query Match          100.0%;  Score 101;  DB 1;  Length 101;
  Best Local Similarity 100.0%;  Pred. No. 2.72e-148;
  Matches   101;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 LGKFSQTCYNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKWQPSNFIETCRN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 LGKFSQTCYNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKWQPSNFIETCRN 60

Db     61 TQLAGSSELAAECKTRAQQFVSTKINLDDHIANIDGTLKYE 101
          |||||||||||||||||||||||||||||||||||||||||
Qy     61 TQLAGSSELAAECKTRAQQFVSTKINLDDHIANIDGTLKYE 101
```

```
RESULT    2
ID   W67569 standard; Protein; 101 AA.
AC   W67569;
DT   02-MAR-1999  (first entry)
DE   N. ellipsosporum cyanovirin protein.
KW   Antiviral protein; cyanovirin; inhibition; infectivity; replication;
KW   cytopathy; virus; HIV; infection.
OS   Nostoc ellipsosporum.
PN   US5843882-A.
PD   01-DEC-1998.
PF   27-APR-1995; 429965.
PR   27-APR-1995; US-429965.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI   Boyd MR, Gustafson KR, McMahon JB, Shoemaker RH;
DR   WPI; 99-044625/04.
DR   N-PSDB; V34401.
PT   Nostoc ellipsosporum proteins or peptide(s) - with antiviral
PT   activity
PS   Claim 7; Column 25-26; 30pp; English.
CC   This sequence represents an antiviral protein, designated cyanovirin,
CC   from Nostoc ellipsosporum.  The antiviral protein, or peptide of at least
CC   9 amino acid residues, is used to inhibit the infectivity, replication
CC   and cytopathic effects of viruses, especially HIV-1 or HIV-2, in the
CC   treatment or prevention of viral infections.
SQ   Sequence   101 AA;

  Query Match          100.0%;  Score 101;  DB 1;  Length 101;
  Best Local Similarity 100.0%;  Pred. No. 2.72e-148;
  Matches   101;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 LGKFSQTCYNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKWQPSNFIETCRN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 LGKFSQTCYNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKWQPSNFIETCRN 60

Db     61 TQLAGSSELAAECKTRAQQFVSTKINLDDHIANIDGTLKYE 101
          |||||||||||||||||||||||||||||||||||||||||
Qy     61 TQLAGSSELAAECKTRAQQFVSTKINLDDHIANIDGTLKYE 101
```

```
RESULT    3
ID   W06811 standard; Protein; 101 AA.
AC   W06811;
DT   17-MAR-1997  (first entry)
DE   Cyanovirin-N.
KW   Cyanovirin-N; cyanobacterium; antiviral; virucide; HIV-1; HIV-2;
KW   SIV; human immunodeficiency virus; retrovirus; AIDS; therapy.
OS   Nostoc ellipsosporum.
PN   WO9634107-A2.
PD   31-OCT-1996.
PF   26-APR-1996; U05908.
PR   27-APR-1995; US-429965.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI   Boyd MR,  Gustafson KR,  McMahon JB,  Shoemaker RH;
DR   WPI; 96-497638/49.
DR   N-PSDB; T45978.
PT   Antiviral protein from Nostoc ellipsosporum - used for treating or
PT   preventing viral infections, esp. infections caused by retro:viruses
PT   such as HIV
PS   Claim 2; Page 78; 99pp; English.
CC   Cyanovirin-N (W06811) of Nostoc ellipsosporum shows antiviral
CC   activity against immunodeficiency retroviruses, esp. HIV-1, HIV-2
CC   and SIV.  It was detected in aq. extracts of the cyanobacterium
CC   using an HIV-specific bioassay-guided strategy.  It can obtd. by
CC   expression in host (esp. yeast, lactobacilli) cells transformed
CC   with a vector carrying cyanovirin-N sequences (see also T45978-79).
CC   It can also be produced as a conjugate with e.g. a toxin (esp.
CC   Pseudomonas exotoxin) or immunological agent.  It is used to treat
CC   or prevent viral infections, and to prevent the spread of such
CC   infections by treating inanimate objects, ex vivo blood, blood
CC   prods. or tissue.
SQ   Sequence   101 AA;

  Query Match          100.0%;  Score 101;  DB 1;  Length 101;
```

```
  Best Local Similarity 100.0%;  Pred. No. 2.72e-148;
  Matches   101;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      1 LGKFSQTCYNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKWQPSNFIETCRN 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 LGKFSQTCYNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKWQPSNFIETCRN 60

Db     61 TQLAGSSELAAECKTRAQQFVSTKINLDDHIANIDGTLKYE 101
          |||||||||||||||||||||||||||||||||||||||||
Qy     61 TQLAGSSELAAECKTRAQQFVSTKINLDDHIANIDGTLKYE 101


RESULT    4
ID   W80553 standard; Protein; 109 AA.
AC   W80553;
DT   09-DEC-1998  (first entry)
DE   Antiviral protein, cyanovirin-N.
KW   Cyanovirin-N; recombinant; cyanovirin; cyanobacterium; antiviral;
KW   HIV-1.
OS   Nostoc ellipsosporum.
FH   Key             Location/Qualifiers
FT   Peptide         1. .8
FT                   /note= "FLAG octapeptide"
PN   US5821081-A.
PD   13-OCT-1998.
PF   26-APR-1996; 638610.
PR   26-APR-1996; US-638610.
PR   27-APR-1995; US-429965.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI   Boyd MR, Gustafson KR, McMahon JB, Shoemaker RH;
DR   WPI; 98-567657/48.
DR   N-PSDB; V56026.
PT   DNA encoding cyanovirin polypeptide(s) - useful for producing
PT   recombinant polypeptides with antiviral activity
PS   Claim 1; Fig 2; 33pp; English.
CC   This represents an antiviral protein cyanovirin-N. A vector containing
CC   a nucleic acid molecule (V56025 or V56026) can be used to transform a
CC   host cell for the recombinant production of the cyanovirin polypeptides.
CC   The cyanovirins are derived from the cyanobacterium Nostoc ellipsosporum,
CC   and have antiviral activity, e.g. with EC50 values of 0.4-7.6 nM against
CC   various HIV-1 strains and isolates.
SQ   Sequence   109 AA;

  Query Match          100.0%;  Score 101;  DB 1;  Length 109;
  Best Local Similarity 100.0%;  Pred. No. 2.72e-148;
  Matches   101;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      9 LGKFSQTCYNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKWQPSNFIETCRN 68
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 LGKFSQTCYNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKWQPSNFIETCRN 60

Db     69 TQLAGSSELAAECKTRAQQFVSTKINLDDHIANIDGTLKYE 109
          |||||||||||||||||||||||||||||||||||||||||
Qy     61 TQLAGSSELAAECKTRAQQFVSTKINLDDHIANIDGTLKYE 101


RESULT    5
ID   W67570 standard; Protein; 109 AA.
AC   W67570;
DT   02-MAR-1999  (first entry)
DE   FLAG epitope-cyanovirin fusion protein.
KW   Antiviral protein; cyanovirin; inhibition; infectivity; replication;
KW   cytopathy; virus; HIV; infection; epitope.
OS   Synthetic.
OS   Nostoc ellipsosporum.
FH   Key             Location/Qualifiers
FT   Region          1. .8
FT                   /label= FLAG_epitope
FT   Region          10. .109
FT                   /label= cyanovirin_N
PN   US5843882-A.
PD   01-DEC-1998.
PF   27-APR-1995; 429965.
PR   27-APR-1995; US-429965.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI   Boyd MR, Gustafson KR, McMahon JB, Shoemaker RH;
DR   WPI; 99-044625/04.
DR   N-PSDB; V34402.
PT   Nostoc ellipsosporum proteins or peptide(s) - with antiviral
PT   activity
PS   Disclosure; Column 27-28; 30pp; English.
CC   This sequence represents a synthetic fusion protein comprising the
CC   antiviral protein, designated cyanovirin, from Nostoc ellipsosporum
CC   with a FLAG epitope peptide fused at its N-terminus.  The antiviral
CC   protein, or peptide of at least 9 amino acid residues, is used to
CC   inhibit the infectivity, replication and cytopathic effects of viruses,
CC   especially HIV-1 or HIV-2, in the treatment or prevention of viral
CC   infections.
SQ   Sequence   109 AA;

  Query Match          100.0%;  Score 101;  DB 1;  Length 109;
  Best Local Similarity 100.0%;  Pred. No. 2.72e-148;
  Matches   101;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      9 LGKFSQTCYNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKWQPSNFIETCRN 68
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 LGKFSQTCYNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKWQPSNFIETCRN 60

Db     69 TQLAGSSELAAECKTRAQQFVSTKINLDDHIANIDGTLKYE 109
          |||||||||||||||||||||||||||||||||||||||||
Qy     61 TQLAGSSELAAECKTRAQQFVSTKINLDDHIANIDGTLKYE 101


RESULT    6
ID   W06812 standard; Protein; 109 AA.
AC   W06812;
DT   17-MAR-1997  (first entry)
DE   FLAG-cyanovirin-N fusion protein.
KW   Cyanovirin-N; cyanobacterium; antiviral; virucide; HIV-1; HIV-2;
KW   SIV; human immunodeficiency virus; retrovirus; AIDS; therapy.
OS   Chimeric Nostoc ellipsosporum;
OS   Chimeric synthetic.
FH   Key             Location/Qualifiers
FT   peptide         1. .8
FT                   /label= FLAG
FT   protein         9. .109
FT                   /label= Cyanovirin-N
PN   WO9634107-A2.
PD   31-OCT-1996.
PF   26-APR-1996; U05908.
PR   27-APR-1995; US-429965.
PA   (USSH ) US DEPT HEALTH & HUMAN SERVICES.
PI   Boyd MR,  Gustafson KR,  McMahon JB,  Shoemaker RH;
DR   WPI; 96-497638/49.
DR   N-PSDB; T45979.
PT   Antiviral protein from Nostoc ellipsosporum - used for treating or
PT   preventing viral infections, esp. infections caused by retro:viruses
PT   such as HIV
PS   Example 2; Page 79-80; 99pp; English.
CC   A polypeptide (W06812) comprises a FLAG octapeptide fused to the
CC   antiviral cyanovirin-N (see also W06811) of Nostoc ellipsosporum.
CC   It can be produced in transformed host cells utilising the pFLAG-1
CC   vector including a synthetic gene (T45979) and purified using
CC   anti-FLAG antibodies.  Cyanovirin-N shows antiviral activity
CC   against immunodeficiency retroviruses, esp. HIV-1, HIV-2 and SIV.
CC   It can also be produced as a conjugate with e.g. a toxin (esp.
CC   Pseudomonas exotoxin) or immunological agent.  It is used to treat
CC   or prevent viral infections, and to prevent the spread of such
CC   infections by treating inanimate objects, ex vivo blood, blood
CC   prods. or tissue.
SQ   Sequence   109 AA;

  Query Match          100.0%;  Score 101;  DB 1;  Length 109;
  Best Local Similarity 100.0%;  Pred. No. 2.72e-148;
  Matches   101;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Db      9 LGKFSQTCYNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKWQPSNFIETCRN 68
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy      1 LGKFSQTCYNSAIQGSVLTSTCERTNGGYNTSSIDLNSVIENVDGSLKWQPSNFIETCRN 60

Db     69 TQLAGSSELAAECKTRAQQFVSTKINLDDHIANIDGTLKYE 109
          |||||||||||||||||||||||||||||||||||||||||
Qy     61 TQLAGSSELAAECKTRAQQFVSTKINLDDHIANIDGTLKYE 101


RESULT    7
ID   R95970 standard; peptide; 299 AA.
AC   R95970;
DT   19-FEB-1997  (first entry)
DE   Nucleic acid recognition unit #6.
KW   Target binding assembly; nucleic acid recognition unit; NF-kappa-B; SP1;
KW   TATA; human papillomavirus; HPV; HPV E2; human immunodeficiency virus;
KW   HIV LTR; Tat binding unit; probe; assembly sequence; asymmetry sequence;
KW   nuclear localisation signal sequence; human cell; HPV LTR; HIV; bacteria;
KW   virus.
OS   Synthetic.
[redacted]WO9617956-A2.
[redacted]13-JUN-1996.
PF   07-DEC-1995; U15944.
PR   09-DEC-1994; US-353476.
PA   (GENE-) GENE POOL INC.
PI   Weininger AM,  Weininger S;
DR   WPI; 96-287199/29.
PT   Probe nucleic acids, target binding assemblies, etc - for detection
PT   and localisation of specific nucleic acid sequences, esp. HIV and
PT   HPV
PS   Claim 14; Page 92-93; 172pp; English.
CC   R95965-R95993 represent the nucleic acid recognition units (NAR) of
CC   target binding assemblies (TBA) of the invention.  These NARs are
CC   selected from NF-kappa-B, SP1, TATA, human papillomavirus (HPV) E2, HPV
CC   LTR, human immunodeficiency virus (HIV) LTR and Tat binding units.  This
CC   sequence represents a NF-kappa-B NAR.  The TBA is recognised by the
CC   target binding region (TBR) of a probe of the invention.  The probe of
CC   the invention contains a TBR, a booster binding region (BBR), and an
CC   optional support or attachment (OSA).  The TBA contains at least one NAR,
CC   and optionally a linker sequence, an assembly sequence, an asymmetry
CC   sequence, a nuclear localisation signal sequence, and an OSA.  The
CC   assembly sequence and asymmetry sequences are responsible for the folding
CC   and association of the NARs.  The linker sequence is an oligopeptide,
CC   which does not interfere with NAR function, but provides stability and
CC   control over the spacing of the NAR from the rest of the TBA.  The OSA is
CC   an attached support or indicator, or other means of localisation of the
CC   probe.  The probe can be used in a method for detecting or localising a
CC   specific target nucleic acid sequence (TNA).  The method is highly
[redacted]sensitive, and has a high degree of specificity.  The method can be used
CC   for detecting specific nucleic acid sequences, including those found in
CC   human cells, in HIV, HPV, and other nucleic acid containing systems,
CC   including bacteria and viruses.
SQ   Sequence   299 AA;

  Query Match             7.9%;  Score 8;  DB 1;  Length 299;
  Best Local Similarity 100.0%;  Pred. No. 3.39e-01;
  Matches     8;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    181 SSIDLNSV 188
          ||||||||
Qy     32 SSIDLNSV 39


RESULT    8
ID   W20743 standard; protein; 443 AA.
AC   W20743;
DT   16-JUL-1997  (first entry)
DE   H. pylori cytoplasmic protein, 06ee10709orf5.
KW   Cytoplasmic; vaccine; prevention; treatment; infection; identification;
KW   binding compound; bacterium; life cycle; activator; bacteria; inhibitor;
KW   duodenal ulcer disease; chronic gastritis; diagnosis; envelope.
OS   Helicobacter pylori.
PN   WO9640893-A1.
PD   19-DEC-1996.
PF   06-JUN-1996; U09122.
PR   07-JUN-1995; US-487032.
PR   01-APR-1996; US-630405.
PA   (ASTR ) ASTRA AB.
PI   Berglindh OT,  Smith D,  Mellgaerd BL;
DR   WPI; 97-052306/05.
DR   N-PSDB; T67996.
PT   Helicobacter pylori nucleic acid sequences and related
PT   polypeptide(s)  - useful for vaccines to treat or prevent H. pylori
PT   infection, and to detect Helicobacter
PS   Claim 61; Page 1158-1159; 1481pp; English.
CC   The present sequence is a Helicobacter pylori cytoplasmic protein
CC   involved in cofactor metabolism.
CC   The protein may be used in a vaccine to prevent or treat H. pylori
CC   infection or to identify H. pylori polypeptide binding compounds,
CC   useful as potential H. pylori life cycle activators or inhibitors.
CC   The genomic sequence of H. pylori (ATCC 55679) was determined from
CC   overlapping contigs generated by mechanically shearing the bacterial
CC   DNA. The sequences were analysed for ORF of at least 180 nucleotides,
CC   and the predicted coding regions defined by computer evaluation. To
CC   identify likely H. pylori antigens for vaccine development, the amino
CC   acid sequences predicted from various ORF were analysed for significant
CC   homology to other known or exported membrane proteins. Having identified
CC   and determined the sequences of interest, particular regions can be
CC   isolated from H. pylori by PCR amplification for recombinant polypeptide
CC   production, e.g. in E. coli hosts.
SQ   Sequence   443 AA;

  Query Match             6.9%;  Score 7;  DB 1;  Length 443;
  Best Local Similarity 100.0%;  Pred. No. 5.52e+00;
  Matches     7;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    330 LNSVIEN 336
          |||||||
Qy     36 LNSVIEN 42


RESULT    9
ID   W14832 standard; peptide; 13 AA.
AC   W14832;
DT   23-MAY-1997  (first entry)
DE   PDGF-1 oncogene protein residues 1-12 (sic).
KW   Oncogene; monoclonal receptor; antibody; immunoglobulin; ligand;
KW   immunogen; epitope; oncoprotein; detection.
OS   Homo sapiens.
PN   US5030565-A.
PD   09-JUL-1991.
PF   15-FEB-1985; 701954.
PR   17-AUG-1983; US-524084.
PR   17-AUG-1984; WO-U01304.
PR   17-AUG-1984; US-001304.
PR   15-FEB-1985; US-701954.
PR   16-APR-1987; US-039534.
PR   18-APR-1991; US-687710.
PR   17-DEC-1993; US-170649.
PA   (SCRI ) SCRIPPS RES INST.
PI   Lerner RA,  Niman HL;
DR   WPI; 91-222277/30.
PT   Monoclonal receptors to protein, esp. onco-protein ligands - prepd.
PT   using a polypeptide corresp. to a portion of the protein aminoacid
PT   sequence
PS   Disclosure; Page -; 41pp; English.
CC   The sequences given in W14803-32 represent peptides derived from
CC   oncogenes which are bound by the monoclonal receptors of the invention.
CC   The monoclonal receptor molecules are immunoglobulins which bind to
CC   both (a) a protein ligand and (b) a polypeptide having an amino acid
CC   residue sequence containing 7-40 amino acid residues corresponding to a
CC   sequence of a portion of the protein, the receptor molecule having been
CC   raised to an immunogen containing the polypeptide. High yields of
CC   monoclonal receptors can be obtained which bind to or immunoreact with
CC   known predetermined epitopes of protein molecules such as oncoproteins.
CC   The receptors can be used for e.g. detection of oncoprotein ligands or
```

```
CC   in affinity sorbants for binding and purifying oncoprotein ligands.
SQ   Sequence   13 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 13;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      8 AECKTR 13
          ||||||
Qy     71 AECKTR 76


RESULT   10
ID   W79087 standard; Protein; 48 AA.
AC   W79087;
DT   11-JAN-1999  (first entry)
DE   Human secreted protein bd164_7.
KW   Secreted protein; human; bd164_7.
OS   Homo sapiens.
     WO9841539-A2.
     24-SEP-1998.
     19-MAR-1998; U05474.
PR   18-MAR-1998; US-040963.
PR   19-MAR-1997; US-820493.
PA   (GEMY ) GENETICS INST INC.
PI   Agostino MJ, Jacobs K, LaVallie ER, McCoy JM, Merberg D,
PI   Racie LA, Spaulding V, Treacy M;
DR   WPI; 98-521163/44.
DR   N-PSDB; V61477.
PT   New polynucleotide(s) encoding secreted human proteins - derived
PT   from human foetal kidney, adult testes and adult or foetal brain
PT   cDNA libraries
PS   Claim 9; Page 67; 112pp; English.
CC   This is the amino acid sequence of a novel human secreted protein,
CC   designated bd164_7.  The sequence was deduced from a full-length
CC   cDNA clone (see V61477) obtained from a foetal kidney cDNA
CC   library.  The protein shows some homology to database sequences.
CC   The invention provides cDNA clones (see V61477-87) from human
CC   foetal kidney, adult testis, and adult or foetal brain cDNA
CC   libraries that code for secreted proteins (see W79087-97).  These
CC   clones are deposited as ATCC 98364.  The polynucleotides and
CC   proteins are predicted to have useful biological activities which
CC   would make them suitable for treating, preventing or ameliorating
CC   medical conditions in humans and animals, although no supporting
CC   data is given.  Suggested activities include nutritional, immune
CC   stimulating (e.g. as vaccines) or suppressing, haematopoiesis
CC   regulating, tissue growth, activin/inhibin, chemotactic/chemokinetic,
     haemostatic and thrombolytic, receptor/ligand, antiinflammatory,
     cadherin/tumour invasion suppressor and tumour inhibition
     activities.
SQ   Sequence   48 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 48;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     14 TCRNTQ 19
          ||||||
Qy     57 TCRNTQ 62


RESULT   11
ID   R25443 standard; Protein; 109 AA.
AC   R25443;
DT   13-JAN-1993  (first entry)
DE   PDGF analogue #3.
KW   platelet derived growth factor;  chemotactic; mitogenic; fibroblasts;
KW   wound healing; dermal ulcers; lacerations; abrasions;
KW   surgical wounds; burns; defined culture media; v-sis protein; p28sis;
KW   neoplasm; cancer; tumour; inhibit atherosclerosis.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   region          26. .34
FT                   /note= "amino acids from corresponding A chain
FT                   sequence"
PN   US5128321-A.
PD   07-JUL-1992.
PF   08-AUG-1988; 230190.
PR   12-OCT-1984; US-660496.
PR   25-FEB-1985; US-705175.
PR   13-AUG-1986; US-896485.
PR   15-DEC-1986; US-941970.
PA   (ZYMO ) ZYMOGENETICS INC.
PI   Kelly JD, Murray MJ;
DR   WPI; 92-249442/30.
PT   New platelet-derived growth factor analogues - comprising protein
PT   homo:dimers having two polypeptide chains with sequences the same
PT   as the A- and opt. B chain obtd. in large amts. used to enhance the
PT   healing of wounds
PS   Disclosure; Fig 3A; 25pp; English.
CC   This sequence represents a new platelet-derived growth factor
CC   analogue chain . The analogue functions as a homodimer of two chains,
CC   each of the chains being a mosaic of amino acid sequences identical
CC   to portions of the A- or B-chains of PDGF, the protein being
CC   chemotactic or mitogenic for fibroblasts. The homodimers have the
CC   same biological activity as PDGF, but can be obtained in large
CC   amts. by recombinant DNA techniques. These PDGF analogues can be
CC   used in a compsn. to enhance the healing of wounds eg. dermal
CC   ulcers, superficial wounds and lacerations, abrasions, surgical
CC   wounds and some burns. The PDGF analogues can also be used as
CC   components of defined, serum-free culture media. They can also be
CC   used in elucidating the putative role of the vsis protein p28sis
CC   in the neoplastic process or for developing inhibitors or
CC   designing specific therapeutic approaches which prevent or
CC   interfere with the in vivo activity of PDGF in individuals with
CC   atherosclerosis.  This sequence is a B chain coding sequence
CC   which has been altered by replacing amino acids 26-34 with the
CC   corresponding A chain amino acids.
SQ   Sequence   109 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 109;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     14 AECKTR 19
          ||||||
Qy     71 AECKTR 76


RESULT   12
ID   R42810 standard; Protein; 109 AA.
AC   R42810;
DT   04-MAY-1994  (first entry)
DE   Mature human PDGF-B.
KW   Platelet-derived growth factor; monomeric; binding; inhibition;
KW   stenosis; restenosis; antiproliferative; invasive cardiovascular;
KW   procedures.
OS   Homo sapiens.
PN   WO9320204-A.
PD   14-OCT-1993.
PF   26-MAR-1993; U02612.
PR   30-MAR-1992; US-860711.
PA   (SCHE ) SCHERING CORP.
PI   Cable MB, Hesson TE, Mannarino AF;
DR   WPI; 93-336912/42.
PT   Monomeric platelet-derived growth factor - useful for preventing
PT   stenosis or restenosis following invasive cardiovascular procedures
PS   Disclosure; Page 27; 41pp; English.
CC   The sequence is that of the mature form of monomeric human
CC   platelet-derived growth factor (PDGF-B). It is identical to a
CC   subsequence of the c-sis protein (Ratner et al. 1985).
SQ   Sequence   109 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 109;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
Db     14 AECKTR 19
          ||||||
Qy     71 AECKTR 76


RESULT   13
ID   P81030 standard; Protein; 109 AA.
AC   P81030;
DT   29-MAR-1992  (first entry)
DE   Sequence of mature B-chain of platelet derived growth factor
DE   (PDGF).
KW   Wound healing; fibroblast; mammalian cell culture.
OS   Homo sapiens.
PN   EP-259632-A.
PD   16-MAR-1988.
PF   11-AUG-1987; 111591.
PR   12-OCT-1984; US-660496.
PR   25-FEB-1985; US-705175.
PR   13-AUG-1986; US-896485.
PR   15-DEC-1986; US-942484.
PR   15-DEC-1986; US-942161.
PR   15-DEC-1986; US-941970.
PA   (ZYMO-) ZYMOGENETICS INC.
PI   Murray MJ, Kelly JD;
DR   WPI; 88-072136/11.
PT   Platelet-derived growth factor analogues - prepd. using DNA
PT   constructs contg. a transcriptional promoter followed by A and/or
PT   B chain coding sequences
PS   Disclosure; Fig 9; 88pp; English.
CC   Computer analysis of a partial AA SQ of platelet derived growth
CC   factor (PDGF) has revealed extensive homology with the gene product,
CC   p28sis, of simian sarcoma virus (SSV). Human PDGF A-chain and
CC   B-chain cDNAs are isolated from a human cDNA library using the v-sis
CC   gene or a fragment as a hybridisation probe. The inventors claim
CC   rDNA constructs for the prodn. of PDGF analogues in eucaryotic cells,
CC   esp. yeast. The PDGF analogues comprise PDGF A-chains and B-chains
CC   as monomers and as heterodimers. A protein having two polypeptide
CC   chains, one of said chains being a mosaic of AA SQs substantially
CC   identical to portions of the A- or B-chains of PDGF, the second of
CC   said chains being substantially homologous to the B-chain of PDGF
CC   is also claimed.
SQ   Sequence   109 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 109;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     14 AECKTR 19
          ||||||
Qy     71 AECKTR 76


RESULT   14
ID   R20967 standard; peptide; 109 AA.
AC   R20967;
DT   20-MAY-1992  (first entry)
DE   Sequence of B-chain of native human platelet derived growth factor
DE   (PDGF).
KW   Wound healing; proteolysis-resistant; platelet derived growth factor.
OS   Homo sapiens.
PN   WO9201716-A.
PD   06-FEB-1992.
PF   22-JUL-1991; U05183.
PR   23-JUL-1990; US-557219.
PA   (ZYMO-) ZYMOGENETICS INC.
PI   Murray MJ;
DR   WPI; 92-064894/08.
PT   New protease resistant platelet derived growth factor-like
PT   proteins - comprising mutated B chain linked to native A or B
PT   chain, useful in wound healing
PS   Disclosure; Fig 1; 43pp; English.
CC   The inventors claim a PDGF B-chain having a substitution or deletion
CC   at posn. 27, 28, 32, 79, 80 or 81. Also claimed are a PDGF-like
CC   protein comprising an altered B-chain bonded to PDGF A-chain or
CC   another B-chain, and a DNA molecule encoding the B-chain. The
CC   mutation of the B-chain pref. comprises substitution of Arg 28 or 32
CC   with one of Pro, Ser, Trp, Gln, His, Met or Asn. The PDGF-like
CC   proteins are resistent to proteolysis and therefore are homogeneous.
SQ   Sequence   109 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 109;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     14 AECKTR 19
          ||||||
Qy     71 AECKTR 76


RESULT   15
ID   R87515 standard; peptide; 109 AA.
AC   R87515;
DT   21-JUN-1996  (first entry)
DE   PDGF mosaic peptide B26.
KW   Platelet-derived growth factor; PDGF; unglycosylated protein; burn;
KW   wound healing; non-healing dermal ulcer; superficial wound; laceration;
KW   abrasion; surgical wound; advanced age; diabetes; cancer; steroid;
KW   anti-inflammatory drug; anticoagulant; homodimer.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   region          1. .25
FT                   /note= "PDGF B chain fragment"
FT   region          26. .34
FT                   /note= "PDGF A chain residues 26-34"
FT   region          35. .109
FT                   /note= "PDGF B chain fragment"
PN   US5474982-A.
PD   12-DEC-1995.
PF   13-AUG-1986; 896485.
PR   13-AUG-1986; US-896485.
PR   15-DEC-1986; US-941970.
PR   08-AUG-1988; US-230190.
PR   30-JUN-1992; US-906544.
PR   12-JUL-1994; US-273779.
PA   (ZYMO ) ZYMOGENETICS INC.
PI   Kelly JD,  Murray MJ;
DR   WPI; 96-039529/04.
PT   New platelet-derived growth factor analogues - useful for promoting
PT   wound healing
PS   Example 2; Fig 3a; 25pp; English.
CC   R87513-R87521 are representitive peptides of the invention.  These
CC   sequences are all mosaics of the platelet-derived growth factor (PDGF) A
CC   chain.  This sequence has residues 26-34 of the B chain sequence
CC   replaced by the corresponding A chain residues.  These sequences are
CC   used to form the novel unglycosylated protein homodimers of the
CC   invention.  These homodimers can be used in compositions for enhancing
CC   the wound healing process.  The protein is expected to accelerate the
CC   healing process in a broad spectrum of wound conditions including
CC   disruptions of the dermal layer of the skin (such as non-healing dermal
CC   ulcers, superficial wounds and lacerations, abrasions, surgical wounds,
CC   and some burns).  These sequences are especially used in enhancing the
CC   wound-healing process in conditions where the normal wound healing
CC   process is suppressed or inhibited, such as advanced age, diabetes,
CC   cancer, and treatment with anti-inflammatory drugs, steroids or
CC   anticoagulants.
SQ   Sequence   109 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 109;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     14 AECKTR 19
          ||||||
Qy     71 AECKTR 76
```

```
RESULT   16
ID   R25673 standard; peptide; 109 AA.
AC   R25673;
DT   15-JAN-1993  (first entry)
DE   PDGF-B.
KW   c-sis; pharmaceutical compositions; wound healing.
OS   Homo sapiens.
PN   EP-495638-A.
PD   22-JUL-1992.
PF   15-JAN-1992; 300330.
PR   16-JAN-1991; US-641345.
PA   (SCHE ) SCHERING CORP.
PI   Alexander DM, Cable MB, Dalie BL, Narula SK;
DR   WPI; 92-243474/30.
PT   Expression of mature human platelet derived growth factor-B -
PT   e.g. using plasmid pTacBlq in E. coli
PS   Disclosure; Page 12; 19pp; English.
CC   This sequence is an unglycosylated, biologically active, mature
CC   human platelet derived growth factor-B (PDGF-B). This sequence is
CC   identical to the sequence of c-sis. This sequence can be used for
CC   any medical condition susceptible to treatment by known PGDF's.
CC   Pharmaceutical compositions for such uses comprise an effective
CC   amount of the PDGF-B and a carrier. It can be used for wound
CC   healing and to treat skin damaged by cuts, abrasions, sun, wind, etc.
SQ   Sequence   109 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 109;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     14 AECKTR 19
          ||||||
Qy     71 AECKTR 76


RESULT   17
ID   R15643 standard; Protein; 109 AA.
AC   R15643;
DT   11-FEB-1992  (first entry)
DE   PDGF B-chain polypeptide.
KW   Platelet derived growth factor; burns; ulcers; radiation wounds.
OS   Homo sapiens.
PN   WO9116335-A.
PD   31-OCT-1991.
PF   25-APR-1991; U02766.
PR   26-APR-1990; US-515474.
PA   (CHIR-) CHIRON CORP.
PI   Nascimento CG, Calderoncacia MD;
DR   WPI; 91-339745/46.
DR   N-PSDB; Q15184.
PT   New hyper-glycosylated platelet derived growth factor - obtained
PT   from yeast, used to treat burns, radiation wounds and ulcers such as
PT   decubitus and cutaneous ulcers
PS   Disclosure; Fig 3; 79pp; English.
CC   The amino acid sequence encodes human platelet derived growth factor
CC   (PDGF) B-chain polypeptide. This forms part of the hyper-glycosylated
CC   PDGF which has as much as three times the mitogenic activity as
CC   that of less glycosylated PDGF. This can be applied topically to
CC   wounds such as cutaneous, dermal, mucosal or epithelial wounds in
CC   humans or animals. It can be used to treat burns, radiation wounds
CC   and ulcers such as decubiti and cutaneous ulcers caused by vascular,
CC   haematologic and metabolic diseases, infections or neoplasms. See
CC   also R15644 and R15645.
SQ   Sequence   109 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 109;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     14 AECKTR 19
          ||||||
Qy     71 AECKTR 76


RESULT   18
ID   R87517 standard; peptide; 110 AA.
AC   R87517;
DT   21-JUN-1996  (first entry)
DE   PDGF mosaic peptide B/A54.
KW   Platelet-derived growth factor; PDGF; unglycosylated protein; burn;
KW   wound healing; non-healing dermal ulcer; superficial wound; laceration;
KW   abrasion; surgical wound; advanced age; diabetes; cancer; steroid;
KW   anti-inflammatory drug; anticoagulant; homodimer.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   region          1. .54
FT                   /note= "PDGF B chain fragment"
FT   region          55. .104
FT                   /note= "PDGF A chain residues 49-104"
PN   US5474982-A.
PD   12-DEC-1995.
PF   13-AUG-1986; 896485.
PR   13-AUG-1986; US-896485.
PR   15-DEC-1986; US-941970.
PR   08-AUG-1988; US-230190.
PR   30-JUN-1992; US-906544.
PR   12-JUL-1994; US-273779.
PA   (ZYMO ) ZYMOGENETICS INC.
PI   Kelly JD,  Murray MJ;
DR   WPI; 96-039529/04.
PT   New platelet-derived growth factor analogues - useful for promoting
PT   wound healing
PS   Example 2; Fig 3b; 25pp; English.
CC   R87513-R87521 are representitive peptides of the invention.  These
CC   sequences are all mosaics of the platelet-derived growth factor (PDGF) A
CC   chain.  This sequence has residues 1-54 of the B chain sequence, followed
CC   by residues 49-104 of the A chain sequence.  These sequences are used to
CC   form the novel unglycosylated protein homodimers of the invention.  These
CC   homodimers can be used in compositions for enhancing the wound healing
CC   process.  The protein is expected to accelerate the healing process in a
CC   broad spectrum of wound conditions including disruptions of the dermal
CC   layer of the skin (such as non-healing dermal ulcers, superficial wounds
CC   and lacerations, abrasions, surgical wounds, and some burns).  These
CC   sequences are especially used in enhancing the wound-healing process in
CC   conditions where the normal wound healing process is suppressed or
CC   inhibited, such as advanced age, diabetes, cancer, and treatment with
CC   anti-inflammatory drugs, steroids or anticoagulants.
SQ   Sequence   110 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 110;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     14 AECKTR 19
          ||||||
Qy     71 AECKTR 76


RESULT   19
ID   R25445 standard; Protein; 110 AA.
AC   R25445;
DT   13-JAN-1993  (first entry)
DE   PDGF analogue #5.
KW   platelet derived growth factor;  chemotactic; mitogenic; fibroblasts;
KW   wound healing; dermal ulcers; lacerations; abrasions;
KW   surgical wounds; burns; defined culture media; v-sis protein; p28sis;
KW   neoplasm; cancer; tumour; inhibit atherosclerosis.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   region          1. .54
FT                   /note= "B chain amino acids"
FT   region          55. .110
FT                   /note= "A chain amino acids"
PN   US5128321-A.
PD   07-JUL-1992.
```

```
PF   08-AUG-1988; 230190.
PR   12-OCT-1984; US-660496.
PR   25-FEB-1985; US-705175.
PR   13-AUG-1986; US-896485.
PR   15-DEC-1986; US-941970.
PA   (ZYMO ) ZYMOGENETICS INC.
PI   Kelly JD, Murray MJ;
DR   WPI; 92-249442/30.
PT   New platelet-derived growth factor analogues - comprising protein
PT   homo:dimers having two polypeptide chains with sequences the same
PT   as the A- and opt. B chain obtd. in large amts. used to enhance the
PT   healing of wounds
PS   Disclosure; Fig 3A; 25pp; English.
CC   This sequence represents a new platelet-derived growth factor
CC   analogue chain . The analogue functions as a homodimer of two chains,
CC   each of the chains being a mosaic of amino acid sequences identical
CC   to portions of the A- or B-chains of PDGF, the protein being
CC   chemotactic or mitogenic for fibroblasts. The homodimers have the
CC   same biological activity as PDGF, but can be obtained in large
CC   amts. by recombinant DNA techniques. These PDGF analogues can be
CC   used in a compsn. to enhance the healing of wounds eg. dermal
CC   ulcers, superficial wounds and lacerations, abrasions, surgical
CC   wounds and some burns. The PDGF analogues can also be used as
CC   components of defined, serum-free culture media. They can also be
CC   used in elucidating the putative role of the vsis protein p28sis
CC   in the neoplastic process or for developing inhibitors or
CC   designing specific therapeutic approaches which prevent or
CC   interfere with the in vivo activity of PDGF in individuals with
CC   atherosclerosis.  This sequence is B chain amino acids 1-54
CC   which have been joined to A chain amino acids 49-104.
SQ   Sequence   110 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 110;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     14 AECKTR 19
          ||||||
Qy     71 AECKTR 76


RESULT   20
ID   R25444 standard; Protein; 110 AA.
AC   R25444;
DT   13-JAN-1993  (first entry)
DE   PDGF analogue #4.
KW   platelet derived growth factor;  chemotactic; mitogenic; fibroblasts;
KW   wound healing; dermal ulcers; lacerations; abrasions;
KW   surgical wounds; burns; defined culture media; v-sis protein; p28sis;
KW   neoplasm; cancer; tumour; inhibit atherosclerosis.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   region          1. .53
FT                   /note= "B chain amino acids"
FT   region          54. .110
FT                   /note= "A chain amino acids"
PN   US5128321-A.
PD   07-JUL-1992.
PF   08-AUG-1988; 230190.
PR   12-OCT-1984; US-660496.
PR   25-FEB-1985; US-705175.
PR   13-AUG-1986; US-896485.
PR   15-DEC-1986; US-941970.
PA   (ZYMO ) ZYMOGENETICS INC.
PI   Kelly JD, Murray MJ;
DR   WPI; 92-249442/30.
PT   New platelet-derived growth factor analogues - comprising protein
PT   homo:dimers having two polypeptide chains with sequences the same
PT   as the A- and opt. B chain obtd. in large amts. used to enhance the
PT   healing of wounds
PS   Disclosure; Fig 3A; 25pp; English.
CC   This sequence represents a new platelet-derived growth factor
CC   analogue chain . The analogue functions as a homodimer of two chains,
CC   each of the chains being a mosaic of amino acid sequences identical
CC   to portions of the A- or B-chains of PDGF, the protein being
CC   chemotactic or mitogenic for fibroblasts. The homodimers have the
CC   same biological activity as PDGF, but can be obtained in large
CC   amts. by recombinant DNA techniques. These PDGF analogues can be
CC   used in a compsn. to enhance the healing of wounds eg. dermal
CC   ulcers, superficial wounds and lacerations, abrasions, surgical
CC   wounds and some burns. The PDGF analogues can also be used as
CC   components of defined, serum-free culture media. They can also be
CC   used in elucidating the putative role of the vsis protein p28sis
CC   in the neoplastic process or for developing inhibitors or
CC   designing specific therapeutic approaches which prevent or
CC   interfere with the in vivo activity of PDGF in individuals with
CC   atherosclerosis.  This sequence is B chain amino acids 1-53
CC   which have been joined to A chain amino acids 48-104.
SQ   Sequence   110 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 110;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     14 AECKTR 19
          ||||||
Qy     71 AECKTR 76


RESULT   21
ID   R68619 standard; Protein; 110 AA.
AC   R68619;
DT   01-SEP-1995  (first entry)
DE   Mature human PDGF-B.
KW   Platelet-derived growth factor beta; PDGF-B; precursor; multimer;
KW   multicistronic expression unit; recombinant protein production.
OS   Homo sapiens.
PN   DE4319708-A.
PD   15-DEC-1994.
PF   10-JUN-1993; 319708.
PR   10-JUN-1993; DE-319708.
PA   (BEIE ) BEIERSDORF AG.
PA   (GBFB ) GBF GES BIOTECH FORSCHUNG GMBH.
PI   Eichner W,  McCarthy JEG,  Schneppe B;
DR   WPI; 95-023815/04.
DR   N-PSDB; Q80480.
PT   Recombinant multimeric protein prodn. in bacteria - transfected
PT   with vector contg. new multi-cistronic expression unit, providing
PT   high level prodn. of pharmaceutical and cosmetic products
PS   Claim 11; Page 30; 42pp; German.
CC   Plasmid pMVW-2 contains the cDNA of the human PDGF-B gene, but with
CC   the 5'-translated region of the precursor being incomplete. The
CC   BamHI/NcoI fragment of pMVW-2 was used as the basis for mutagenesis
CC   to isolate the sequence coding for mature PDGF-B with a Met residue
CC   at the N-terminus (R68619). The mutagenised coding sequence (Q80480)
CC   was used in the construction of multicistronic expression units for
CC   recombinant expression of the PDGF-A/B dimer in bacteria.
SQ   Sequence   110 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 110;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     15 AECKTR 20
          ||||||
Qy     71 AECKTR 76


RESULT   22
ID   R87516 standard; peptide; 110 AA.
AC   R87516;
DT   21-JUN-1996  (first entry)
DE   PDGF mosaic peptide B/A53.
KW   Platelet-derived growth factor; PDGF; unglycosylated protein; burn;
KW   wound healing; non-healing dermal ulcer; superficial wound; laceration;
KW   abrasion; surgical wound; advanced age; diabetes; cancer; steroid;
```

```
KW   anti-inflammatory drug; anticoagulant; homodimer.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   region          1. .53
FT                   /note= "PDGF B chain fragment"
FT   region          54. .110
FT                   /note= "PDGF A chain fragment"
PN   US5474982-A.
PD   12-DEC-1995.
PF   13-AUG-1986; 896485.
PR   13-AUG-1986; US-896485.
PR   15-DEC-1986; US-941970.
PR   08-AUG-1988; US-230190.
PR   30-JUN-1992; US-906544.
PR   12-JUL-1994; US-273779.
PA   (ZYMO ) ZYMOGENETICS INC.
PI   Kelly JD, Murray MJ;
DR   WPI; 96-039529/04.
PT   New platelet-derived growth factor analogues - useful for promoting
PT   wound healing
PS   Example 2; Fig 3a; 25pp; English.
CC   R87513-R87521 are representitive peptides of the invention. These
CC   sequences are all mosaics of the platelet-derived growth factor (PDGF) A
CC   chain. This sequence has residues 1-53 of the B chain sequence, followed
CC   by residues 48-104 of the A chain sequence. These sequences are used to
CC   form the novel unglycosylated protein homodimers of the invention. These
CC   homodimers can be used in compositions for enhancing the wound healing
CC   process. The protein is expected to accelerate the healing process in a
CC   broad spectrum of wound conditions including disruptions of the dermal
CC   layer of the skin (such as non-healing dermal ulcers, superficial wounds
CC   and lacerations, abrasions, surgical wounds, and some burns). These
CC   sequences are especially used in enhancing the wound-healing process in
CC   conditions where the normal wound healing process is suppressed or
CC   inhibited, such as advanced age, diabetes, cancer, and treatment with
CC   anti-inflammatory drugs, steroids or anticoagulants.
SQ   Sequence   110 AA;

  Query Match            5.9%;  Score 6;  DB 1;  Length 110;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     14 AECKTR 19
          ||||||
Qy     71 AECKTR 76


RESULT   23
     P80163 standard; protein; 114 AA.
     P80163;
     16-NOV-1990  (first entry)
DE   Biosynthetic multifunctional protein.
KW   Biosynthetic multifunctional protein; biosynthetic antibody binding site;
KW   protein trailer; platelet derived growth factor-beta.
PN   WO8809344-A.
PD   01-DEC-1988.
PF   19-MAY-1988; U01737.
PR   21-MAY-1987; US-052800.
PA   (CREA-) Creative Biomolecules Inc.
PI   Huston JS, Oppermann H;
DR   WPI; 88-353928/49.
DR   N-PSDB; N80189.
PT   Recombinant multifunctional protein - having antibody binding site and a
PT   sequence for biological activity, ion sequestering or binding to a
PT   solid support.
PS   Disclosure; 115pp; English.
CC   The sequence is a biosynthetic multifunctional protein including a
CC   biosynthetic antibody binding site and a platelet derived growth
CC   factor-beta protein trailer linked via a spacer sequence.
SQ   Sequence   114 AA;

  Query Match            5.9%;  Score 6;  DB 1;  Length 114;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     19 AECKTR 24
          ||||||
Qy     71 AECKTR 76


RESULT   24
ID   R04020 standard; peptide; 114 AA.
AC   R04020;
DT   18-MAR-1993  (first entry)
DE   PDGF-BB monomer unit.
KW   Platelet derived growth factor; fibroblast; smooth muscle cell;
KW   monomer.
OS   Synthetic.
PN   DE3834079-A.
PD   12-APR-1990.
PF   06-OCT-1988; 834079.
PR   06-OCT-1988; DE-834079.
PA   (GBFB ) GBF GES BIOTECH FORSCH.
PI   Hoppe J, Eichner W, Weich H;
DR   WPI; 90-116760/16.
PT   New homo:dimeric form of platelet derived growth factor -
PT   stimulates growth of fibroblasts and smooth muscle cells
PS   Claim 1; Page 5; 9pp; German.
CC   A biologically active PDGF-BB  contains monomer units having the
CC   sequence below.  Both monomeric units can have (independent of the
CC   other) amino acid deletions or substitutions or include extra amino
CC   acids.  Esp. the C-termini of both monomeric units may have
CC   deletions or extensions of up to 14 amino acids.
CC   PDGF-BB actively stimulates growth of fibroblast and smooth muscle
CC   cells and can be used therapeutically.
SQ   Sequence   114 AA;

  Query Match            5.9%;  Score 6;  DB 1;  Length 114;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     14 AECKTR 19
          ||||||
Qy     71 AECKTR 76


RESULT   25
ID   R26047 standard; Protein; 114 AA.
AC   R26047;
DT   27-JAN-1993  (first entry)
DE   PDGF-B7.
KW   Dimer; analogue; PDGF-Bt; PDGF-A.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   misc_difference 32
FT                   /label= R32P
PN   WO9213073-A.
PD   06-AUG-1992.
PF   24-JAN-1992; G00141.
PR   25-JAN-1991; GB-001645.
PA   (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
PA   (PFIZ ) PFIZER LTD.
PI   Brown D, Clements JM, Cook AL, Craig S, Edwards RM;
DR   WPI; 92-284658/34.
DR   N-PSDB; Q27208.
PT   Protease-resistant platelet-derived growth factor-B analogues -
PT   for treating atherosclerosis and wounds, having aminoacid(s)
PT   replaced by residue which reduces or prevents cleavage after
PT   specific aminoacid
PS   Claim 7; Page 58; 71pp; English.
CC   Platelet-derived growth factor (PDGF)-B7 is a PDGF-Bt derivative.
CC   The Arg at coding position 32 has been altered to Pro, which is the
CC   same residue which is found in PDGF-A.  This PDGF-B analogue can be
CC   used in the production of a dimer.  This dimer can be useful in
CC   veterinary or human medicine for preparing a medicament for promoting
CC   wound healing and/or treating atherosclerosis.  The analogue, when
CC   dimerised has a biological activity equivalent to a naturally-occuring
```

```
CC   PDGF-B dimer.
SQ   Sequence   114 AA;

  Query Match            5.9%;  Score 6;  DB 1;  Length 114;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     19 AECKTR 24
          ||||||
Qy     71 AECKTR 76


RESULT   26
ID   R41523 standard; Protein; 120 AA.
AC   R41523;
DT   04-MAR-1994  (first entry)
DE   rPDGF refolded B119 homodimer.
KW   Platelet derived-growth factor; recombinantly-derived;
KW   refolded B-chain; enhance; enhancement; proliferation; activation;
KW   epithelial cells; keratinocytes; cornea; accelerate; wound healing;
KW   re-innervation; corneal epithelium; laser surgery; diabetes.
OS   Homo sapiens.
PN   WO9316719-A.
PD   02-SEP-1993.
PF   24-FEB-1993; U01731.
PR   26-FEB-1992; US-842306.
PA   (ALLR ) ALLERGAN INC.
PI   Nicolson MA, Stern ME, Wheeler LA;
DR   WPI; 93-288121/36.
DR   N-PSDB; Q48409.
PT   Accelerating corneal wound healing - using platelet-derived growth
PT   factor to improve quality of healing and accelerate re-innervation
PS   Example 1; Fig 14; 60pp; English.
CC   The sequence is that of a recombinantly-derived platelet-derived
CC   growth factor refolded B-chain homodimer of 119 amino acids.
CC   It may be used to enhance proliferation and activation of
CC   epithelial cells and/or keratocytes of the cornea to accelerate
CC   wound healing, improve the quality of wound healing (by decreasing
CC   epithelial sloughing and/or decreasing scar formation) and
CC   accelerate re-innervation of the corneal epithelium. The wounds may
CC   be as a result of laser surgery or diabetes.
SQ   Sequence   120 AA;

  Query Match            5.9%;  Score 6;  DB 1;  Length 120;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     15 AECKTR 20
          ||||||
Qy     71 AECKTR 76


RESULT   27
ID   R67261 standard; Protein; 120 AA.
AC   R67261;
DT   23-JUN-1995  (first entry)
DE   PDGF-B119.
KW   PDGF-B119; platelet-derived growth factor; tissue growth factor;
KW   collateral circulation; anastomosis; ischemia.
OS   Homo sapiens.
PN   WO9425056-A.
PD   10-NOV-1994.
PF   29-APR-1994; U04762.
PR   29-APR-1993; US-055062.
PA   (UNIW ) UNIV WASHINGTON.
PI   Khouri RK;
DR   WPI; 95-006213/01.
DR   N-PSDB; Q75291.
PT   Use of platelet-derived growth factor to improve collateral
PT   circulation  - administered to tissues at risk of ischaemia
PS   Disclosure; Fig. 1; 59pp; English.
CC   PDGF-B119 corresponds to the first 119 amino acids of the c-sis
CC   encoded PDGF-B precursor protein.  Encoding DNA, obtained from
CC   c-sis or by modification of v-sis can be used to prepare
CC   recombinant PDGF-BB119 homodimer in Escherichia coli.
SQ   Sequence   120 AA;

  Query Match            5.9%;  Score 6;  DB 1;  Length 120;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     15 AECKTR 20
          ||||||
Qy     71 AECKTR 76


RESULT   28
ID   R12879 standard; Protein; 120 AA.
AC   R12879;
DT   18-SEP-1991  (first entry)
DE   rPDGF-B119 from pCFM1156.
KW   Recombinant platelet-derived growth factor; B chain; PDGF-B109;
KW   v-sis-gene; SSV.
OS   Simian sarcoma virus.
PN   WO9108761-A.
PD   27-JUN-1991.
PF   17-DEC-1990; U07451.
PR   19-DEC-1989; US-454794.
PR   13-DEC-1990; US-624451.
PA   (AMGE-) AMGEN INC.
PI   Thomason AR;
DR   WPI; 91-207869/28.
DR   N-PSDB; Q12487.
PT   Platelet-derived growth factor B chain analogues - used to treat
PT   wounds, to aid in healing of bone, cartridge, tendons, ligaments
PT   and epithelium
PS   Disclosure; Fig 33(a-b); 59pp; English.
CC   The v-sis-gene (the oncogen present in SSV), encoding a protein
CC   more than 90% homologous with the PDGF-B chain, is subjected to
CC   in vitro mutagenesis to change codons 101 and 107, then cut back
CC   (to amino acid 24) with BglII.  The portion removed is replaced by
CC   a synthetic sequence corresp. to c-sis PP (i.e. providing
CC   conversion of amino acids 6-7 to the form present in PDGF).  A similar
CC   cut-back is used at the C-terminus to change amino acids 114 and to
CC   introduce a stop codon at amino acid position 120.
CC   The final construct is then incorporated into an expression vector
CC   and the recombinant prod. esp. used to transform E. coli.
CC   The recombinant prod. is isolated, refolded and purified.
CC   See also WO9108762.
SQ   Sequence   120 AA;

  Query Match            5.9%;  Score 6;  DB 1;  Length 120;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     15 AECKTR 20
          ||||||
Qy     71 AECKTR 76


RESULT   29
ID   R60614 standard; Protein; 120 AA.
AC   R60614;
DT   19-MAY-1995  (first entry)
DE   Human PDGF-B 119 subunit.
KW   Human platelet derived growth factor; PDGF-B subunit; v-sis gene;
KW   PDGF-BB fusion protein; wound healing.
OS   Homo sapiens.
PN   EP-618227-A.
PD   05-OCT-1994.
PF   31-MAR-1994; 105075.
PR   01-APR-1993; US-041635.
PA   (AMGE-) AMGEN INC.
PI   Thomason AR;
DR   WPI; 94-304405/38.
DR   N-PSDB; Q71557.
```

```
PT   New multimeric polypeptide comprising fused subunits of natural
PT   protein.  - and related DNA and transformed cells, esp. dimers of
PT   platelet derived growth factor, useful for stimulating healing of
PT   wounds.
PS   Example 1; Page 16; 30pp; English.
CC   The PDGF-B 119 subunit coding sequence was obtained by mutagenesis
CC   of the human v-sis sequence; Ile 101 and Ala 107 codons were
CC   replaced by Thr and Pro codons, respectively. Then the 5'-end was
CC   replaced by a synthetic fragment which changed amino acids 6 and 7
CC   to the corresp. amino acids found in the human form. Finally Thr
CC   114 was replaced by Gly and a stop codon was introduced at position
CC   120. The subunit was incorporated into a single polypeptide with a
CC   PDGF-B 109 subunit, pref. separated by a peptide linker to produce
CC   preferred fusion polypeptide of the invention. The fusion
CC   polypeptide is useful for treating wounds.
SQ   Sequence   120 AA;

  Query Match            5.9%;  Score 6;  DB 1;  Length 120;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches    6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     15 AECKTR 20
          ||||||
Qy     71 AECKTR 76


RESULT   30
ID   R12881 standard; Protein; 129 AA.
AC   R12881;
DT   24-SEP-1991  (first entry)
DE   rPDGF B109 [Ser43, Ser52, Ser53, Ser99].
KW   Platelet derived growth factor; disulphide bond; inclusion bodies;
KW   PDGF B.
OS   Synthetic.
PN   WO9108762-A.
PD   27-JUN-1991.
PF   17-DEC-1990; U07460.
PR   15-DEC-1989; US-451485.
PR   12-DEC-1990; US-623671.
PA   (AMGE-) AMGEN INC.
PI   Lyons DE, Thomason AR;
DR   WPI; 91-207870/28.
DR   N-PSDB; Q12489.
PT   Refolding reduced platelet derived growth factor - by conversion
PT   to mixed di:sulphide with blocking agent, converting to active
PT   form and de-blocking esp. for inclusion bodies in E coli
PS   Disclosure; Fig 10(A-B); 79pp; English.
CC   The starting material was the v-sis gene, with the four amino acid
CC   differences between the v-sis protein and human PDGF B (i.e.
CC   residues 6, 7, 100 and 107) being changed to the human form of PDGF B.
CC   These changes involved mutagenesis of amino acids 100 and 107, and
CC   replacement of the 5'end of the gene with a synthetic DNA fragment
CC   changing amino acids 6 and 7.  The synthetic DNA fragment also
CC   provided a translation initiating ATG codon preceding the Ser1 codon, as
CC   well as an upstream XbaI site for ligation into the desired E. coli
CC   expression vector.  The resulting DNA was used as a template for further
CC   mutagenesis resactions to introduce four cysteine to serine changes and
CC   a stop codon at positon 110.  The four Cys residues which were
CC   changed are those residues at positons 43, 52, 53 and 99.
CC   The prod. was refolded into a biologically active monomer without
CC   interference from incorrect disulphide bond formation.
CC   See also WO9108761.
SQ   Sequence   129 AA;

  Query Match            5.9%;  Score 6;  DB 1;  Length 129;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches    6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     15 AECKTR 20
          ||||||
Qy     71 AECKTR 76


RESULT   31
ID   W64063 standard; Protein; 146 AA.
AC   W64063;
DT   06-OCT-1998  (first entry)
DE   Chimeric rhPDGF-B expression cassette protein.
KW   Platelet derived growth factor; PDGF-B; fusion protein; secretion;
KW   expression; yeast; human.
OS   Homo sapiens.
OS   Saccharomyces cerevisiae.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   Peptide         1. .37
FT                   /note= "truncated yeast alpha factor leader"
FT   Protein         38. .146
FT                   /label= rhPDGF-B
FT                   /note= "human platelet derived growth factor B"
PN   WO9826080-A1.
PD   18-JUN-1998.
PF   12-DEC-1997; U22647.
PR   13-DEC-1996; US-032720.
PA   (CHIR ) CHIRON CORP.
PI   Merryweather JP, Tekamp-Olson P;
DR   WPI; 98-348534/30.
DR   N-PSDB; V44136.
PT   Nucleotide sequence for expression of heterologous proteins in yeast
PT   - useful to produce heterologous mammalian proteins in biologically
PT   active, mature form, e.g. human platelet-derived growth factor
PS   Example 2; Page 49-50; 81pp; English.
CC   This sequence represents a novel fusion protein cassette constructed from
CC   the Saccharomyces cerevisiae alpha leader peptide and human platelet
CC   derived growth factor type B (PDGF-B). This protein is used in a method
CC   which allows the expression of heterologous mammalian proteins and their
CC   secretion in the biologically active, mature form, by transforming yeast
CC   host cells with the vector and expressing the proteins. It is
CC   particularly useful to produce mammalian proteins whose assumption of a
CC   native conformation is facilitated by the presence of a native
CC   propeptide sequence in the precursor polypeptide. Expression and
CC   secretion of proteins in biologically active, mature form is possible,
CC   whilst previous methods have often encountered problems of
CC   post-translational processing leading to, e.g. production of inactive
CC   forms or insufficient amounts of protein.
SQ   Sequence   146 AA;

  Query Match            5.9%;  Score 6;  DB 1;  Length 146;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches    6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     51 AECKTR 56
          ||||||
Qy     71 AECKTR 76


RESULT   32
ID   W64062 standard; Protein; 146 AA.
AC   W64062;
DT   06-OCT-1998  (first entry)
DE   Chimeric rhPDGF-B protein.
KW   Platelet derived growth factor; PDGF-B; fusion protein; secretion;
KW   expression; yeast; human.
OS   Homo sapiens.
OS   Saccharomyces cerevisiae.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   Peptide         1. .37
FT                   /label= leader
FT                   /note= "yeast alpha factor leader peptide"
FT   Protein         38. .146
FT                   /label= PDGF-B
FT                   /note= "human platelet derived growth factor type B"
PN   WO9826080-A1.
PD   18-JUN-1998.
PF   12-DEC-1997; U22647.
PR   13-DEC-1996; US-032720.
```

```
PA   (CHIR ) CHIRON CORP.
PI   Merryweather JP, Tekamp-Olson P;
DR   WPI; 98-348534/30.
DR   N-PSDB; V44114.
PT   Nucleotide sequence for expression of heterologous proteins in yeast
PT   - useful to produce heterologous mammalian proteins in biologically
PT   active, mature form, e.g. human platelet-derived growth factor
PS   Example 1; Page 38; 81pp; English.
CC   This sequence represents a novel fusion protein constructed from the
CC   Saccharomyces cerevisiae alpha leader peptide and human platelet
CC   derived growth factor type B (PDGF-B). This protein is used in a method
CC   which allows the expression of heterologous mammalian proteins and their
CC   secretion in the biologically active, mature form, by transforming yeast
CC   host cells with the vector and expressing the proteins. It is
CC   particularly useful to produce mammalian proteins whose assumption of a
CC   native conformation is facilitated by the presence of a native
CC   propeptide sequence in the precursor polypeptide. Expression and
CC   secretion of proteins in biologically active, mature form is possible,
CC   whilst previous methods have often encountered problems of
CC   post-translational processing leading to, e.g. production of inactive
CC   forms or insufficient amounts of protein.
SQ   Sequence   146 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 146;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      51 AECKTR 56
           ||||||
Qy      71 AECKTR 76


RESULT   33
ID   R26045 standard; Protein; 161 AA.
AC   R26045;
DT   27-JAN-1993  (first entry)
DE   Synthetic PDGF-B.
KW   S. cerevisiae; E. coli; dimer; analogue.
OS   Synthetic.
PN   WO9213073-A.
PD   06-AUG-1992.
PF   24-JAN-1992; G00141.
PR   25-JAN-1991; GB-001645.
PA   (BRBI-) BRITISH BIO-TECHNOLOGY LTD.
PA   (PFIZ ) PFIZER LTD.
PI   Brown D, Clements JM, Cook AL, Craig S, Edwards RM;
DR   WPI; 92-284658/34.
DR   N-PSDB; Q27204.
PT   Protease-resistant platelet-derived growth factor-B analogues -
PT   for treating atherosclerosis and wounds, having aminoacid(s)
PT   replaced by residue which reduces or prevents cleavage after
PT   specific aminoacid
PS   Disclosure; Fig 3; 71pp; English.
CC   The gene encoding this sequence is a synthetic gene which was designed
CC   with codons selected which are favoured by either S. cerevisiae or E.
CC   coli and is thus suitable for expression in either organism.  This
CC   platelet-derived growth factor (PDGF) analogue can be used in the
CC   production of a dimer.  This dimer can be useful in veterinary or
CC   human medicine for preparing a medicament for promoting wound healing
CC   and/or treating atherosclerosis.  The analogue, when dimerised has a
CC   biological activity equivalent to a naturally-occuring PDGF-B dimer.
SQ   Sequence   161 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 161;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      15 AECKTR 20
           ||||||
Qy      71 AECKTR 76


RESULT   34
ID   R87507 standard; Protein; 175 AA.
AC   R87507;
DT   15-OCT-1995  (first entry)
DE   Caenorhabditis elegans her-1 protein.
KW   Nematode; her-1; hermaphroditization; biological control agent;
KW   male differentiation.
OS   Caenorhabditis elegans.
FH   Key             Location/Qualifiers
FT   protein         1. .175
FT                   /label= her-1 protein
PN   US5472871-A.
PD   05-DEC-1995.
PF   28-FEB-1992; 844294.
PR   28-FEB-1992; US-844294.
PR   09-FEB-1994; US-194180.
PA   (COLS ) UNIV COLORADO FOUND INC.
PI   Perry MD,  Trent C,  Wood WB;
DR   WPI; 96-029813/03.
DR   N-PSDB; T06745, T06746.
PT   Nematode her-1 gene - encodes a protein which induces male
PT   differentiation, used for controlling nematodes
PS   Claim 1; Fig.3; 29pp; English.
CC   This her-1 protein from C. elegans prevents female development in
CC   nematodes and activates the male development pathway.  It can be
CC   used for genetically engineering plants, bacteria and endophytes
CC   to control nematode infestation in plants.  The her-1 protein can
CC   be used for the control of nematode infestation in plants, animals
CC   and humans.  This protein may be expressed recombinantly in E. coli
CC   using an expression vector.
SQ   Sequence   175 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 175;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      81 KINLDD 86
           ||||||
Qy      84 KINLDD 89


RESULT   35
ID   R50012 standard; Protein; 190 AA.
AC   R50012;
DT   29-SEP-1994  (first entry)
DE   Truncated Platelet derived growth factor PDGF-B190 mutant.
KW   Platelet-Derived Growth Factor; heterodimer; PDGF-AB;
KW   recombinant protein production; truncated PDGF-B chain; secretable;
KW   bicistronic vector system; PCR; mutagenesis.
OS   Synthetic.
FH   Key             Location/Qualifiers
FT   peptide         1. .81
FT                   /label= pre-peptide
FT   protein         82. .190
FT                   /label= mature_PDGF-B
PN   WO9405785-A.
PD   17-MAR-1994.
PF   26-AUG-1993; E02294.
PR   27-AUG-1992; DE-228458.
PA   (BEIE ) BEIERSDORF AG.
PA   (GBFB ) GBF GES BIOTECH FORSCHUNG GMBH.
PI   Achterberg V,  Dirks W,  Dorschner A,  Eichner W,  Hauser H;
PI   Meyer-Ingold W,  Mielke H,  Wirth M,  Doerschner A;
DR   WPI; 94-101190/12.
DR   N-PSDB; Q58733.
PT   New multicistronic expression units - for producing equimolar
PT   amts. of polypeptide(s) in mammalian cells as hosts
PS   Claim 12; Page 68-69; 109pp; German.
CC   In order to increase the amount of recombinantly produced PDGF-B
CC   that is secreted from a host, a truncated version of the protein
CC   is produced. A stop codon is introduced at amino acid position 191
CC   with the result that the mature protein is still produced but the
CC   additional C-terminal sequence responsible for retaining PDGF-B at
CC   the cell surface is removed. The coding sequence is preferred for
```

```
CC   use in a bicistronic expression system for the recombinant
CC   production of a PDGF-AB heterodimer.
SQ   Sequence   190 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 190;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     95 AECKTR 100
          ||||||
Qy     71 AECKTR 76


RESULT   36
ID   R60615 standard; Protein; 190 AA.
AC   R60615;
DT   19-MAY-1995  (first entry)
DE   Human PDGF-B 109 subunit precursor.
KW   Human platelet derived growth factor; PDGF-B subunit; v-sis gene;
KW   PDGF-BB fusion protein; wound healing.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   protein         82. .190
FT                   /label= PDGF-B109
FT   region          1. .81
FT                   /label= pre-pro-region
PN   EP-618227-A.
PD   05-OCT-1994.
PF   31-MAR-1994; 105075.
PR   01-APR-1993; US-041635.
PA   (AMGE-) AMGEN INC.
PI   Thomason AR;
DR   WPI; 94-304405/38.
DR   N-PSDB; Q71558.
PT   New multimeric polypeptide comprising fused subunits of natural
PT   protein.  - and related DNA and transformed cells, esp. dimers of
PT   platelet derived growth factor, useful for stimulating healing of
PT   wounds.
PS   Example 2; Page 17-18; 30pp; English.
CC   The PDGF-B 109 subunit coding sequence was obtained as composite
CC   sequence of fragments from human PDGF-B pre-pro region, human v-sis
CC   gene and synthetic fragments. The mature 109 amino acid subunit was
CC   incorporated into a single polypeptide with a PDGF-B 119 subunit,
CC   pref. separated by a peptide linker to produce a preferred fusion
CC   polypeptide of the invention. The fusion polypeptide is useful for
CC   treating wounds.
SQ   Sequence   190 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 190;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     95 AECKTR 100
          ||||||
Qy     71 AECKTR 76


RESULT   37
ID   R05615 standard; protein; 198 AA.
AC   R05615;
DT   13-AUG-1990  (first entry)
DE   BIV vtf gene product.
KW   BIV 127; BIV 106; vaccine; AIDS; ds.
OS   Bovine Immunodeficiency Virus.
PN   US7408815-A.
PD   30-JAN-1990.
PF   18-SEP-1989; 408815.
PR   18-SEP-1989; US-408815.
PA   (USSH) US Dept of Health & Human.
PI   Gonda M;
DR   WPI; 90-099193/13.
DR   N-PSDB; Q05056 and Q03642.
PT   Cloned bovine immuno-deficiency like virus -
PT   used for preparing vaccines or making antigens, antibodies or
PT   DNA probes for diagnostic tests.
PS   Disclosure; Fig 6; 65pp; English.
CC   The sequence is useful in manufacturing antigens, and antibodies
CC   for diagnosis, treatment and prophylaxis of the BIV induced disease.
SQ   Sequence   198 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 198;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    141 SVLTST 146
          ||||||
Qy     16 SVLTST 21


RESULT   38
ID   W53971 standard; peptide; 202 AA.
AC   W53971;
DT   19-AUG-1998  (first entry)
DE   Crinipellis scabella wild-type cellulase.
KW   Cellulase; variant; cellulolytic enzyme; endoglucanase V; EGV; paper;
KW   protein engineering; detergent; fabric softener; pulp; debarking;
KW   defibration; fibre modification; degradation.
OS   Crinipellis scabella.
PN   WO9812307-A1.
PD   26-MAR-1998.
PF   17-SEP-1997; DK0393.
PR   17-SEP-1996; DK-001013.
PA   (NOVO ) NOVO-NORDISK AS.
PI   Andersen KV, Christiansen L, Damgaard B, Schulein M;
DR   WPI; 98-217251/19.
PT   Cellulase enzyme variants - having amino acid changes which improve
PT   properties e.g. activity, sensitivity to surfactants, pH optimum or
PT   stability
PS   Disclosure; Page 7-11; 115pp; English.
CC   The present invention describes a cellulase enzyme variant comprising a
CC   catalytic core domain exhibiting cellulolytic activity which is derived
CC   from a naturally occurring parental cellulase by amino acid residue
CC   substitution, insertion and/or deletion; and (with cellulase numbering):
CC   (a) at position 5 has an Ala, Ser, or Thr residue; (b) at position 8 has
CC   a Phe or a Tyr residue; (c) at position 9 has a Phe, Trp, or Tyr residue;
CC   (d) at position 10 has a Asp residue; (e) at position 121 has a Asp
CC   residue. The present sequence represents the Crinipellis scabella wild-
CC   type cellulase used as a basis for the new cellulase enzyme variants. The
CC   new cellulase enzymes can be used in e.g. detergent or fabric softener
CC   compositions, for bio-polishing of new fabrics, for promoting a stone-
CC   washed look to cellulosic containing fabric, for pulp and paper
CC   applications, e.g. for debarking, defibration, fibre modification,
CC   drainage improvement, inter fibre bonding or for degradation of plant
CC   material e.g. for improving feed value. The cellulase variants have
CC   improved properties with respect to e.g. catalytic activity, altered
CC   sensitivity to anionic tensides, pH optimum or activity profile or
CC   stability.
SQ   Sequence   202 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 202;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     76 LAGSSE 81
          ||||||
Qy     63 LAGSSE 68


RESULT   39
ID   W53970 standard; peptide; 204 AA.
AC   W53970;
DT   19-AUG-1998  (first entry)
DE   Myceliophthora thermophila wild-type cellulase.
KW   Cellulase; variant; cellulolytic enzyme; endoglucanase V; EGV; paper;
KW   protein engineering; detergent; fabric softener; pulp; debarking;
KW   defibration; fibre modification; degradation.
```

OS Myceliophthora thermophila.
PN WO9812307-A1.
PD 26-MAR-1998.
PF 17-SEP-1997; DK0393.
PR 17-SEP-1996; DK-001013.
PA (NOVO) NOVO-NORDISK AS.
PI Andersen KV, Christiansen L, Damgaard B, Schulein M;
DR WPI; 98-217251/19.
PT Cellulase enzyme variants - having amino acid changes which improve
PT properties e.g. activity, sensitivity to surfactants, pH optimum or
PT stability
PS Disclosure; Page 7-11; 115pp; English.
CC The present invention describes a cellulase enzyme variant comprising a
CC catalytic core domain exhibiting cellulolytic activity which is derived
CC from a naturally occurring parental cellulase by amino acid residue
CC substitution, insertion and/or deletion; and (with cellulase numbering):
CC (a) at position 5 has an Ala, Ser, or Thr residue; (b) at position 8 has
CC a Phe or a Tyr residue; (c) at position 9 has a Phe, Trp, or Tyr residue;
CC (d) at position 10 has a Asp residue; (e) at position 121 has a Asp
CC residue. The present sequence represents the Myceliophthora thermophila
CC wild-type cellulase used as a basis for the new cellulase enzyme
CC variants. The new cellulase enzymes can be used in e.g. detergent or
CC fabric softener compositions, for bio-polishing of new fabrics, for
CC promoting a stone-washed look to cellulosic containing fabric, for pulp
CC and paper applications, e.g. for debarking, defibration, fibre
CC modification, drainage improvement, inter fibre bonding or for
CC degradation of plant material e.g. for improving feed value. The
CC cellulase variants have improved properties with respect to e.g.
CC catalytic activity, altered sensitivity to anionic tensides, pH optimum
CC or activity profile or stability.
SQ Sequence 204 AA;

Query Match 5.9%; Score 6; DB 1; Length 204;
Best Local Similarity 100.0%; Pred. No. 7.69e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 77 LAGSSE 82
||||||
Qy 63 LAGSSE 68

RESULT 40
ID W64064 standard; Protein; 205 AA.
AC W64064;
DT 06-OCT-1998 (first entry)
DE Chimeric rhPDGF-B protein with human pro-sequence.
KW Platelet derived growth factor; PDGF-B; fusion protein; secretion;
KW expression; yeast; human.
OS Homo sapiens.
OS Saccharomyces cerevisiae.
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1. .35
FT /label= signal
FT /note= "yeast alpha-factor signal/leader sequence"
FT Peptide 36. .96
FT /label= transit
FT /note= "human PDGF propeptide"
FT Protein 97. .615
FT /label= PDGF-B
FT /note= "human platelet derived growth factor type B"
PN WO9826080-A1.
PD 18-JUN-1998.
PF 12-DEC-1997; U22647.
PR 13-DEC-1996; US-032720.
PA (CHIR) CHIRON CORP.
PI Merryweather JP, Tekamp-Olson P;
DR WPI; 98-348534/30.
DR N-PSDB; V44137.
PT Nucleotide sequence for expression of heterologous proteins in yeast
PT - useful to produce heterologous mammalian proteins in biologically
PT active, mature form, e.g. human platelet-derived growth factor
PS Claim 10; Page 52; 81pp; English.
CC This sequence represents a novel fusion protein constructed from the
CC Saccharomyces cerevisiae alpha leader peptide and human platelet derived
CC growth factor type B (PDGF-B) and also contains a human propeptide
CC sequence. This protein is used in a method which allows the expression of
CC heterologous mammalian proteins and their secretion in the biologically
CC active, mature form, by transforming yeast host cells with the vector and
CC expressing the proteins. It is particularly useful to produce mammalian
CC proteins whose assumption of a native conformation is facilitated by the
CC presence of a native propeptide sequence in the precursor polypeptide.
CC Expression and secretion of proteins in biologically active, mature form
CC is possible, whilst previous methods have often encountered problems of
CC post-translational processing leading to, e.g. production of inactive
CC forms or insufficient amounts of protein.
SQ Sequence 205 AA;

Query Match 5.9%; Score 6; DB 1; Length 205;
Best Local Similarity 100.0%; Pred. No. 7.69e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 110 AECKTR 115
||||||
Qy 71 AECKTR 76

RESULT 41
ID W64066 standard; Protein; 205 AA.
AC W64066;
DT 06-OCT-1998 (first entry)
DE Chimeric rhPDGF-B protein with human pro-sequence and ADH/GAP promoter.
KW Platelet derived growth factor; PDGF-B; fusion protein; secretion;
KW expression; yeast; human.
OS Homo sapiens.
OS Saccharomyces cerevisiae.
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1. .37
FT /label= signal
FT /note= "yeast truncated alpha factor leader peptide"
FT Peptide 38. .96
FT /label= propeptide
FT /note= "human PDGF propeptide"
FT Protein 97. .205
FT /label= rhPDGF-B
FT /note= "human platelet derived growth factor B"
PN WO9826080-A1.
PD 18-JUN-1998.
PF 12-DEC-1997; U22647.
PR 13-DEC-1996; US-032720.
PA (CHIR) CHIRON CORP.
PI Merryweather JP, Tekamp-Olson P;
DR WPI; 98-348534/30.
DR N-PSDB; V44145.
PT Nucleotide sequence for expression of heterologous proteins in yeast
PT - useful to produce heterologous mammalian proteins in biologically
PT active, mature form, e.g. human platelet-derived growth factor
PS Example 3; Page 61-62; 81pp; English.
CC This sequence represents a novel fusion protein constructed from the
CC Saccharomyces cerevisiae alpha leader peptide and human platelet derived
CC growth factor type B (PDGF-B) and also contains a human propeptide
CC sequence and an ADH/GAP promoter. This protein is used in a method which
CC allows the expression of heterologous mammalian proteins and their
CC secretion in the biologically active, mature form, by transforming yeast
CC host cells with the vector and expressing the proteins. It is
CC particularly useful to produce mammalian proteins whose assumption of a
CC native conformation is facilitated by the presence of a native propeptide
CC sequence in the precursor polypeptide. Expression and secretion of
CC proteins in biologically active, mature form is possible, whilst previous
CC methods have often encountered problems of post-translational processing
CC leading to, e.g. production of inactive forms or insufficient amounts of
CC protein.
SQ Sequence 205 AA;

Query Match 5.9%; Score 6; DB 1; Length 205;

```
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches    6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    110 AECKTR 115
          ||||||
Qy     71 AECKTR 76


RESULT   42
ID   R68617 standard; Protein; 216 AA.
AC   R68617;
DT   31-AUG-1995  (first entry)
DE   Human PDGF-B precursor sequence (N-terminally truncated).
KW   Platelet-derived growth factor beta; PDGF-B; precursor; multimer;
KW   multicistronic expression unit; recombinant protein production.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   region          1. .56
FT                   /label= pre-region
                     /note= "truncated"
     protein         57. .216
                     /label= mature_PDGF-B
PN   DE4319708-A.
PD   15-DEC-1994.
PF   10-JUN-1993; 319708.
PR   10-JUN-1993; DE-319708.
PA   (BEIE ) BEIERSDORF AG.
PA   (GBFB ) GBF GES BIOTECH FORSCHUNG GMBH.
PI   Eichner W,  McCarthy JEG,  Schneppe B;
DR   WPI; 95-023815/04.
DR   N-PSDB; Q80473.
PT   Recombinant multimeric protein prodn. in bacteria - transfected
PT   with vector contg. new multi-cistronic expression unit, providing
PT   high level prodn. of pharmaceutical and cosmetic products
PS   Example 2.1.2; Page 23-24; 42pp; German.
CC   Plasmid pMVW-2 contains the cDNA of the human PDGF-B gene, but with
CC   the 5'-translated region of the precursor being incomplete (see
CC   Q80473, which encodes R68617). The BamHI/NcoI fragment of pMVW-2 was
CC   used as the basis for mutagenesis to isolate the sequence coding for
CC   mature PDGF-B with a Met residue at the N-terminus. The mutagenised
CC   coding sequence was used in the construction of multicistronic
CC   expression units for recombinant expression of the PDGF-A/B dimer
CC   in bacteria.
SQ   Sequence   216 AA;

  Query Match            5.9%;  Score 6;  DB 1;  Length 216;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches    6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     70 AECKTR 75
          ||||||
Qy     71 AECKTR 76


RESULT   43
ID   R63470 standard; Protein; 220 AA.
AC   R63470;
DT   28-JUN-1995  (first entry)
DE   Recombinant platelet derived growth factor B exons 2-6.
KW   Recombinant platelet derived growth factor B; rPDGF B; exons 2-6;
KW   wound healing; antibody production.
OS   Synthetic.
PN   EP-622456-A.
PD   02-NOV-1994.
PF   10-MAR-1988; 302116.
PR   13-MAR-1987; US-025344.
PR   19-FEB-1988; US-152045.
PA   (AMGE-) AMGEN INC.
PI   Nicolson MA,  Thomason AR;
DR   WPI; 94-334642/42.
DR   N-PSDB; Q78189.
PT   New vectors expressing platelet-derived growth factor B
PT   epitope(s)  - and contg. c-sis or v-sis genes portions; are used
PT   to promote wound healing and produce antibodies
PS   Example 1; Fig 4; 49pp; English.
CC   Q78189 exons 2-6 of recombinant platelet derived growth factor B
CC   (rPDGF B) encodes R63470. rPDGF B may be used to promote wound healing,
CC   or as an antigen to prepare antisera or monoclonal antibodies specific
CC   for epitopes on the B chain of PDGF.
SQ   Sequence   220 AA;

  Query Match            5.9%;  Score 6;  DB 1;  Length 220;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches    6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     74 AECKTR 79
          ||||||
Qy     71 AECKTR 76


RESULT   44
ID   R40965 standard; Protein; 220 AA.
AC   R40965;
DT   25-FEB-1994  (first entry)
DE   rPDGF B.
KW   Monoclonal antibody; specific; epitope; B-Chain; PDGF;
KW   platelet-derived growth factor; wound healing; binding.
OS   Homo sapiens.
PN   EP-559234-A.
PD   08-SEP-1993.
PF   10-MAR-1988; 302116.
PR   13-MAR-1987; US-025344.
PR   19-FEB-1988; US-152045.
PA   (AMGE-) AMGEN.
PI   Nicolson MA,  Thomason AR;
DR   WPI; 93-282094/36.
DR   N-PSDB; Q48656.
PT   New monoclonal antibody for prepn. of poly:peptide for wound
PT   healing - is specific for epitope in B-chain of platelet-derived
PT   growth factor
PS   Example 1; Fig 4; 39pp; English.
CC   Example 1 describes the subcloning and analysis of C-sis and V-sis
CC   genes encoding rPDGF B chain.
CC   Clone U2-OS561 contains the c-sis gene encoding human PDGF B chain.
CC   Exons 2-6 of this clone were subcloned and sequenced. They encode
CC   the mature rPDGF B protein. The exon sequences are the same as those
CC   published for a c-cis gene isolated from a human fetal liver chromosomal
CC   library. (Josephs et al, Science 223: 487 (1984))
SQ   Sequence   220 AA;

  Query Match            5.9%;  Score 6;  DB 1;  Length 220;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches    6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     74 AECKTR 79
          ||||||
Qy     71 AECKTR 76


RESULT   45
ID   W04925 standard; Protein; 225 AA.
AC   W04925;
DT   20-MAY-1997  (first entry)
DE   Cellulytic enzyme #1 of the invention.
KW   Cellulytic enzyme; endoglucanase; hydrolysis; cellulose; microorganism;
KW   plant cellulase; catalytic region; textile; backstaining; bio-polishing;
KW   stone-washing; cellulosic fabric; colour clarification; defibration;
KW   cell wall degradation; paper pulp; debarking; fibre modification;
KW   enzymatic de-inking; drainage improvement.
OS   Myceliophthora thermophila.
PN   WO9629397-A1.
PD   26-SEP-1996.
PF   18-MAR-1996; DK0105.
PR   17-MAR-1995; DK-000272.
PR   08-AUG-1995; DK-000887.
PR   08-AUG-1995; DK-000885.
```

PR 08-AUG-1995; DK-000888.
PR 08-AUG-1995; DK-000886.
PR 12-FEB-1996; DK-000137.
PA (NOVO) NOVO-NORDISK AS.
PI Andersen LN, Ihara M, Kauppinen MS, Lange L, Lassen SF;
PI Nielsen RI, Schuelein M, Takagi S;
DR WPI; 96-443173/44.
DR N-PSDB; T39047.
PT New endo:glucanase enzyme preparations - contg. conserved catalytic
PT regions, useful for treating fabrics, textiles, plant material or
PT paper pulp
PS Claim 66; Page 140-142; 316pp; English.
CC W04925-W04932 represent the enzymes of the invention. These enzymes
CC possess cellulytic (particularly endoglucanase) activity. Cellulytic
CC enzymes are involved in the the hydrolysis of cellulose, and are
CC synthesised by a large number of microorganisms and plants. The enzymes
CC of the invention containing the conserved catalytic regions (such as
CC W04913) exhibit improved performance, e.g. 50 times higher performance,
CC compared to multiple domain enzymes. The enzymes can be used for the
CC treatment of fabrics or textiles, preferably for preventing backstaining,
CC for bio-polishing or for stone-washing cellulosic fabric. They can also
CC be used to provide colour clarification for laundry. The enzymes can also
CC be used for the degradation or modification of plant material, such as
CC cell walls. They can also be used in the treatment of paper pulp
CC preferably for debarking, defibration, fibre modification, enzymatic
CC de-inking or drainage improvement.
SQ Sequence 225 AA;

Query Match 5.9%; Score 6; DB 1; Length 225;
Best Local Similarity 100.0%; Pred. No. 7.69e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 98 LAGSSE 103
||||||
Qy 63 LAGSSE 68

RESULT 46
ID P60215 standard; Protein; 226 AA.
AC P60215;
DT 12-AUG-1991 (first entry)
DE Sequence encoded by the portion of the v-sis gene encoding a protein
DE homologous to the B-chain of platelet derived growth factor (PDGF).
KW Mitogen; mesenchyma; fibroblast proliferation; smooth muscle cell;
KW atherosclerosis therapy.
OS Simian sarcoma virus.
PN EP-177957-A.
PD 16-APR-1986.
PF 10-OCT-1985; 012852.
PR 12-OCT-1984; US-660496.
PR 25-FEB-1985; US-705175.
PR 13-AUG-1986; US-896485.
PR 15-DEC-1986; US-941970.
PR 15-DEC-1986; US-942484.
PA (ZYMO-) ZYMOGENETICS INC.
PI Murray MJ, Kelly JD;
DR WPI; 86-101617/16.
DR N-PSDB; N60147.
PT Prodn. of platelet derived growth factor analogues - useful as
PT mitogenic agents and obtd. by recombinant DNA procedures for good
PT purity and high yields
PS Disclosure; Fig. 1B; 60pp; English.
CC A DNA construct contg. the portion of the v-sis gene encoding a
CC protein substantially homologous to the B chain of PDGF is claimed.
CC Also claimed are biologically active PDGF analogues obtained using
CC the construct.
SQ Sequence 226 AA;

Query Match 5.9%; Score 6; DB 1; Length 226;
Best Local Similarity 100.0%; Pred. No. 7.69e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 80 AECKTR 85
||||||
Qy 71 AECKTR 76

RESULT 47
ID W04930 standard; Protein; 226 AA.
AC W04930;
DT 20-MAY-1997 (first entry)
DE Cellulytic enzyme #6 of the invention.
KW Cellulytic enzyme; endoglucanase; hydrolysis; cellulose; microorganism;
KW plant cellulase; catalytic region; textile; backstaining; bio-polishing;
KW stone-washing; cellulosic fabric; colour clarification; defibration;
KW cell wall degradation; paper pulp; debarking; fibre modification;
KW enzymatic de-inking; drainage improvement.
OS Crinipellis scabella.
PN WO9629397-A1.
PD 26-SEP-1996.
PF 18-MAR-1996; DK0105.
PR 17-MAR-1995; DK-000272.
PR 08-AUG-1995; DK-000887.
PR 08-AUG-1995; DK-000885.
PR 08-AUG-1995; DK-000888.
PR 08-AUG-1995; DK-000886.
PR 12-FEB-1996; DK-000137.
PA (NOVO) NOVO-NORDISK AS.
PI Andersen LN, Ihara M, Kauppinen MS, Lange L, Lassen SF;
PI Nielsen RI, Schuelein M, Takagi S;
DR WPI; 96-443173/44.
DR N-PSDB; T39052.
PT New endo:glucanase enzyme preparations - contg. conserved catalytic
PT regions, useful for treating fabrics, textiles, plant material or
PT paper pulp
PS Claim 76; Page 169-170; 316pp; English.
CC W04925-W04932 represent the enzymes of the invention. These enzymes
CC possess cellulytic (particularly endoglucanase) activity. Cellulytic
CC enzymes are involved in the the hydrolysis of cellulose, and are
CC synthesised by a large number of microorganisms and plants. The enzymes
CC of the invention containing the conserved catalytic regions (such as
CC W04913) exhibit improved performance, e.g. 50 times higher performance,
CC compared to multiple domain enzymes. The enzymes can be used for the
CC treatment of fabrics or textiles, preferably for preventing backstaining,
CC for bio-polishing or for stone-washing cellulosic fabric. They can also
CC be used to provide colour clarification for laundry. The enzymes can also
CC be used for the degradation or modification of plant material, such as
CC cell walls. They can also be used in the treatment of paper pulp
CC preferably for debarking, defibration, fibre modification, enzymatic
CC de-inking or drainage improvement.
SQ Sequence 226 AA;

Query Match 5.9%; Score 6; DB 1; Length 226;
Best Local Similarity 100.0%; Pred. No. 7.69e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 99 LAGSSE 104
||||||
Qy 63 LAGSSE 68

RESULT 48
ID P81029 standard; Protein; 226 AA.
AC P81029;
DT 29-MAR-1992 (first entry)
DE Sequence encoded by the v-sis gene and some flanking SSV sequences
DE corrected from that published by Devare et al. in 1982.
KW Wound healing; fibroblast; mammalian cell culture.
OS Simian sarcoma virus.
PN EP-259632-A.
PD 16-MAR-1988.
PF 11-AUG-1987; 111591.
PR 12-OCT-1984; US-660496.
PR 25-FEB-1985; US-705175.
PR 13-AUG-1986; US-896485.
PR 15-DEC-1986; US-942484.

PR 15-DEC-1986; US-942161.
PR 15-DEC-1986; US-941970.
PA (ZYMO-) ZYMOGENETICS INC.
PI Murray MJ, Kelly JD;
DR WPI; 88-072136/11.
DR N-PSDB; N81371.
PT Platelet-derived growth factor analogues - prepd. using DNA
PT constructs contg. a transcriptional promoter followed by A and/or
PT B chain coding sequences
PS Example; Fig 1B; 88pp; English.
CC Computer analysis of a partial AA SQ of platelet derived growth
CC factor (PDGF) has revealed extensive homology with the gene product,
CC p28sis, of simian sarcoma virus (SSV). Human PDGF A-chain and
CC B-chain cDNAs are isolated from a human cDNA library using the v-sis
CC gene or a fragment as a hybridisation probe. The inventors claim
CC rDNA constructs for the prodn. of PDGF analogues in eucaryotic cells,
CC esp. yeast. The PDGF analogues comprise PDGF A-chains and B-chains
CC as monomers and as heterodimers. A protein having two polypeptide
CC chains, one of said chains being a mosaic of AA SQs substantially
CC identical to portions of the A- or B-chains of PDGF, the second of
CC said chains being substantially homologous to the B-chain of PDGF
CC is also claimed.
SQ Sequence 226 AA;

Query Match 5.9%; Score 6; DB 1; Length 226;
Best Local Similarity 100.0%; Pred. No. 7.69e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 80 AECKTR 85
||||||
Qy 71 AECKTR 76

RESULT 49
ID R22673 standard; Protein; 226 AA.
AC R22673;
DT 25-NOV-1992 (first entry)
DE v-sis protein p28sis.
KW PDGF analogue; platelet derived growth factor; SSV; neoplasm; cancer;
KW atherosclerosis.
OS Simian Sarcoma Virus.
FH Key Location/Qualifiers
FT misc_difference 36
FT /label= Asp encoded by GAG(Glu)
PN EP-487116-A.
PD 27-MAY-1992.
PF 10-OCT-1985; 121387.
PR 12-OCT-1984; US-660496.
PR 25-FEB-1985; US-705175.
PA (ZYMO) ZYMOGENETICS INC.
PI Kelly JD, Murray MJ;
DR WPI; 92-176732/22.
DR N-PSDB; Q25040.
PT New human platelet derived growth factor analogues - have
PT mitogenic activity for promoting the growth of mammalian cells
PS Claim 3; Fig 1B; 32pp; English.
CC This sequence represents the p28sis gene product of the v-sis gene
CC from simian sarcoma virus (SSV). p28sis is highly homologous to the
CC B chain of human PDGF. The SSV retroviral genome was cloned from
CC SSV-II nonproductively infected normal rat kidney cell, and isolated
CC as a 5.8kb Eco RI fragment - see Q25040.
CC The PDGF growth factor analogue is used to promote the growth of
CC mammalian cells by incubating it with the cells. Recombinantly
CC produced PDGF avoids the risks associated with human blood derived
CC products which may be contaminated with hepatitus virus,
CC cytomegalovirus, or HIV. The ability to produce large amounts of
CC PDGF may be useful tool in elucidating the putative role of p28sis
CC in the neoplastic process. It could also be used to develop
CC inhibitors or for designing specific methods to interfere with the in
CC vivo activity of PDGF in individuals with atherosclerosis.
CC See also Q25040,1
SQ Sequence 226 AA;

Query Match 5.9%; Score 6; DB 1; Length 226;
Best Local Similarity 100.0%; Pred. No. 7.69e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 80 AECKTR 85
||||||
Qy 71 AECKTR 76

RESULT 50
ID R63472 standard; protein; 241 AA.
AC R63472;
DT 28-JUN-1995 (first entry)
DE Recombinant platelet derived growth factor B cv-sis.
KW Recombinant platelet derived growth factor B cv-sis; rPDGF Bcv-sis;
KW wound healing; antibody production.
OS Synthetic.
PN EP-622456-A.
PD 02-NOV-1994.
PF 10-MAR-1988; 302116.
PR 13-MAR-1987; US-025344.
PR 19-FEB-1988; US-152045.
PA (AMGE-) AMGEN INC.
PI Nicolson MA, Thomason AR;
DR WPI; 94-334642/42.
PT New vectors expressing platelet-derived growth factor B
PT epitope(s) - and contg. c-sis or v-sis genes portions; are used
PT to promote wound healing and produce antibodies
PS Claim 10; Fig 10; 49pp; English.
CC R63468 describes the amino acid sequence of recombinant platelet
CC derived growth factor B cv-sis (rPDGF B cv-sis). It may be used to
CC promote wound healing, or as an antigen to prepare antisera or
CC monoclonal antibodies specific for epitopes on the B chain of PDGF.
SQ Sequence 241 AA;

Query Match 5.9%; Score 6; DB 1; Length 241;
Best Local Similarity 100.0%; Pred. No. 7.69e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 95 AECKTR 100
||||||
Qy 71 AECKTR 76

RESULT 51
ID R50009 standard; Protein; 241 AA.
AC R50009;
DT 29-SEP-1994 (first entry)
DE Platelet-derived Growth Factor PDGF-B precursor.
KW Platelet-Derived Growth Factor; heterodimer; PDGF-AB;
KW recombinant protein production; PDGF-B chain;
KW bicistronic vector system.
OS Homo sapiens.
FH Key Location/Qualifiers
FT peptide 1. .81
FT /label= pre-peptide
FT protein 82. .190
FT /label= mature_PDGF-B
FT /note= "mature peptide is encoded by nucleotides
FT 283-609 of Q58725"
PN WO9405785-A.
PD 17-MAR-1994.
PF 26-AUG-1993; E02294.
PR 27-AUG-1992; DE-228458.
PA (BEIE) BEIERSDORF AG.
PA (GBFB) GBF GES BIOTECH FORSCHUNG GMBH.
PI Achterberg V, Dirks W, Dorschner A, Eichner W, Hauser H;
PI Meyer-Ingold W, Mielke H, Wirth M, Doerschner A;
DR WPI; 94-101190/12.
DR N-PSDB; Q58725.
PT New multicistronic expression units - for producing equimolar
PT amts. of polypeptide(s) in mammalian cells as hosts
PS Claim 11; Page 29; 109pp; German.

```
CC   A PDGF-AB heterodimer is recombinantly produced using a bicistronic
CC   expression unit in which a sequence responsible for internal
CC   translation start is located between cistrons coding for the PDGF-B
CC   and PDGF-A chains. The preferred PDGF-B sequence for inclusion in
CC   the bicistronic construct is Q58725 which codes for the precursor
CC   amino acid sequence R50009.
SQ   Sequence   241 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 241;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      95 AECKTR 100
           ||||||
Qy      71 AECKTR 76


RESULT   52
ID   R50002 standard; Protein; 241 AA.
AC   R50002;
DT   29-SEP-1994  (first entry)
DE   Platelet-derived Growth Factor PDGF-B precursor.
KW   Platelet-Derived Growth Factor; heterodimer; PDGF-AB;
KW   recombinant protein production; PDGF-B chain;
KW   bicistronic vector system.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   peptide         1. .81
FT                   /label= pre-peptide
FT   protein         82. .190
FT                   /label= mature_PDGF-B
FT                   /note= "mature peptide is encoded by nucleotides
FT                   283-609 of Q58714"
PN   WO9405786-A.
PD   17-MAR-1994.
PF   26-AUG-1993; E02295.
PR   27-AUG-1992; DE-228457.
PA   (BEIE ) BEIERSDORF AG.
PA   (GBFB ) GBF GES BIOTECH FORSCHUNG GMBH.
PI   Achterberg V,  Dirks W,  Dorschner A,  Eichner W,  Hauser H;
PI   Meyer-Ingold W,  Mielke H,  Wirth M,  Doerschner A;
DR   WPI; 94-101191/12.
DR   N-PSDB; Q58714.
PT   Heterodimer platelet-derived-growth factor (PDGF) prodn. - using
PT   a polycistronic vector system in mammalian host cells for
PT   equimolar prodn of A- and B-chains
PS   Claim 2; Page 40; 64pp; German.
CC   A PDGF-AB heterodimer is recombinantly produced using a bicistronic
CC   expression unit in which a sequence responsible for internal
CC   translation start is located between cistrons coding for the PDGF-B
CC   and PDGF-A chains. The preferred PDGF-B sequence for inclusion in
CC   the bicistronic construct is Q58714 which codes for the precursor
CC   amino acid sequence R50002.
SQ   Sequence   241 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 241;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      95 AECKTR 100
           ||||||
Qy      71 AECKTR 76


RESULT   53
ID   P80596 standard; protein; 241 AA.
AC   P80596;
DT   12-OCT-1990  (first entry)
DE   Recombinant platelet derived growth factor Bc-sis.
KW   REcombinant platelet derived growth factor; wound healing;
KW   fusion peptides;  Bc-sis; antibodies.
OS   Simian sarcoma virus.
PN   EP-282317-A.
PD   14-SEP-1988.
PF   10-MAR-1988; 302116.
PR   19-FEB-1988; US-152045.
PA   (AMGE-) AMGEN INC.
PI   Thomason AR, Nicolson; MA;
DR   WPI; 88-259036/37.
PT   Purified recombinant platelet derived growth factor -
PT   for accelerating wound healing, and fusion polypeptide(s),
PT   encoding DNA sequences and derived antibodies.
PS   Disclosure; p; English.
CC   The peptide has a sufficient part of the structural conformation of
CC   PDGF B, an epitope for binding to a monoclonal antibody spaecific
CC   for an epitope of a B chain PDGF, one or more of the biological
CC   properties of naturally occurring PDGF and has purity of greater than
CC   95%. The cpds. is used in veterinary and human medicine for topical
CC   application to wounds to accelerate healing. They are produced free of
CC   PDGF A, dimers and heterodimers. The antibodies are useful for
CC   affinity purification. See also P80595-P80598 and N81202-N81203.
SQ   Sequence   241 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 241;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      95 AECKTR 100
           ||||||
Qy      71 AECKTR 76


RESULT   54
ID   R38919 standard; Protein; 241 AA.
AC   R38919;
DT   28-OCT-1993  (first entry)
DE   Human Platelet Derived growth Factor B-Chain.
KW   Angiogenesis; wound healing; mitogen; vascular endothelial cells;
KW   Vascular Endothelial Cell Growth Factor; bVEGF-120; PDGF;
KW   chimeric growth factor protein.
OS   Homo sapiens.
PN   US5219739-A.
PD   15-JUN-1993.
PF   27-JUL-1989; 387545.
PR   27-JUL-1989; US-387545.
PR   14-DEC-1989; US-450883.
PR   27-JUL-1990; US-559041.
PA   (SCIO-) SCIOS NOVA INC.
PI   Abraham JA,  Fiddes JC,  Mitchell RL,  Tischer EG;
DR   WPI; 93-205302/25.
DR   N-PSDB; Q44258.
PT   Isolated DNA sequences, expression vectors and transformant cells
PT   - used for large scale prodn. of vascular endothelial cell growth
PT   factor, for treating wounds in which neo-vascularisation is
PT   required
PS   Disclosure; Fig 4b; 40pp; English.
CC   Chimeric growth factor proteins are contemplated by the inventors
CC   (but not claimed). An example is a dimeric protein consisting of
CC   the full-length B-chain subunit of PDGF linked by a disulphide bond
CC   to a full-length vascular endothelial cell growth factor
CC   polypeptide chain. Preparation of such hybrid proteins allows the
CC   properties of the molecule to be tailored such that the hybrid
CC   exhibits a profile of mitogenic activity between VEGF and PDGF. .
SQ   Sequence   241 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 241;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      95 AECKTR 100
           ||||||
Qy      71 AECKTR 76


RESULT   55
ID   R40967 standard; Protein; 241 AA.
```

AC R40967;
DT 25-FEB-1994 (first entry)
DE CV-sis gene product.
KW Monoclonal antibody; specific; epitope; B-Chain; PDGF;
KW platelet-derived growth factor; wound healing; binding.
OS Synthetic.
PN EP-559234-A.
PD 08-SEP-1993.
PF 10-MAR-1988; 302116.
PR 13-MAR-1987; US-025344.
PR 19-FEB-1988; US-152045.
PA (AMGE-) AMGEN.
PI Nicolson MA, Thomason AR;
DR WPI; 93-282094/36.
PT New monoclonal antibody for prepn. of poly:peptide for wound
PT healing - is specific for epitope in B-chain of platelet-derived
PT growth factor
PS Example 2; Fig 10; 39pp; English.
CC pDSCE/cv-sis is a mammalian expression vector in which the mature
CC rPDGF B protein region is encoded by the v-sis gene, but the
CC translation initiation signal and amino terminal region of the rPDGF
CC B pre-protein is encoded by the c-sis gene. The amino
CC acid sequence of the expected protein is given.
SQ Sequence 241 AA;

Query Match 5.9%; Score 6; DB 1; Length 241;
Best Local Similarity 100.0%; Pred. No. 7.69e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 95 AECKTR 100
||||||
Qy 71 AECKTR 76

RESULT 56
ID W64065 standard; Protein; 241 AA.
AC W64065;
DT 06-OCT-1998 (first entry)
DE Chimeric rhPDGF-B protein with human prepro-sequence.
KW Platelet derived growth factor; PDGF-B; fusion protein; secretion;
KW expression; yeast; human.
OS Homo sapiens.
OS Saccharomyces cerevisiae.
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1. .22
FT /label= Prepeptide
FT /note= "human PDGF-B prepeptide"
FT Peptide 23. .81
FT /label= transit
FT /note= "human PDGF propeptide"
FT Protein 82. .190
FT /label= rhPDGF-B
FT /note= "human platelet derived growth factor type B"
FT Peptide 191. .241
FT /label= transit
FT /note= "human PDGF propeptide"
PN WO9826080-A1.
PD 18-JUN-1998.
PF 12-DEC-1997; U22647.
PR 13-DEC-1996; US-032720.
PA (CHIR) CHIRON CORP.
PI Merryweather JP, Tekamp-Olson P;
DR WPI; 98-348534/30.
DR N-PSDB; V44138.
PT Nucleotide sequence for expression of heterologous proteins in yeast
PT - useful to produce heterologous mammalian proteins in biologically
PT active, mature form, e.g. human platelet-derived growth factor
PS Example 3; Page 55-56; 81pp; English.
CC This sequence represents a novel fusion protein constructed from the
CC Saccharomyces cerevisiae alpha leader peptide and human platelet derived
CC growth factor type B (PDGF-B) and also contains two human propeptide
CC sequences. This protein is used in a method which allows the expression
CC of heterologous mammalian proteins and their secretion in the
CC biologically active, mature form, by transforming yeast host cells with
CC the vector and expressing the proteins. It is particularly useful to
CC produce mammalian proteins whose assumption of a native conformation is
CC facilitated by the presence of a native propeptide sequence in the
CC precursor polypeptide. Expression and secretion of proteins in
CC biologically active, mature form is possible, whilst previous methods
CC have often encountered problems of post-translational processing leading
CC to, e.g. production of inactive forms or insufficient amounts of protein.
SQ Sequence 241 AA;

Query Match 5.9%; Score 6; DB 1; Length 241;
Best Local Similarity 100.0%; Pred. No. 7.69e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 95 AECKTR 100
||||||
Qy 71 AECKTR 76

RESULT 57
ID R40964 standard; protein; 241 AA.
AC R40964;
DT 25-FEB-1994 (first entry)
DE PDGF Bc-sis.
KW Monoclonal antibody; specific; epitope; B-Chain; PDGF;
KW platelet-derived growth factor; wound healing; binding.
OS Homo sapiens.
PN EP-559234-A.
PD 08-SEP-1993.
PF 10-MAR-1988; 302116.
PR 13-MAR-1987; US-025344.
PR 19-FEB-1988; US-152045.
PA (AMGE-) AMGEN.
PI Nicolson MA, Thomason AR;
DR WPI; 93-282094/36.
PT New monoclonal antibody for prepn. of poly:peptide for wound
PT healing - is specific for epitope in B-chain of platelet-derived
PT growth factor
PS Example 1; Fig 2; 39pp; English.
CC Example 1 describes the subcloning and analysis of C-sis and V-sis
CC genes encoding rPDGF B chain. Monoclonal antibodies against
CC rPDGF Bv-sis have the affinity to specifically bind to an epitope
CC found in the B chain of PDGF.
SQ Sequence 241 AA;

Query Match 5.9%; Score 6; DB 1; Length 241;
Best Local Similarity 100.0%; Pred. No. 7.69e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 95 AECKTR 100
||||||
Qy 71 AECKTR 76

RESULT 58
ID P80597 standard; protein; 241 AA.
AC P80597;
DT 24-OCT-1990 (first entry)
DE Cv-sis gene encoded Platelet Derived Growth Factor.
KW Recombinant platelet derived growth factor; wound healing; cv-sis gene;
KW rPDGF; antibodies.
OS Homo sapiens.
PN EP-282317-A.
PD 14-SEP-1988.
PF 10-MAR-1988; 302116.
PR 19-FEB-1988; US-152045.
PA (AMGE-) Amgen INC.
PI Thomason AR, Nicolson MA;
DR WPI; 88-259036/37.
PT Purified recombinant platelet derived growth factor - for accelerating
PT wound healing, and fusion polypeptide(s), encoding DNA sequences
PT and derived antibodies.

PS Disclosure; pp; English.
CC The mature rPDGF B protein region is encoded by the v-sis gene, but
CC the translation initiation signal and amino terminal of the rPDGF B
CC pre-protein is encoded by the c-sis gene.
CC The product is useful in human and veterinary medicine for the
CC topical application to wounds to accelerate healing. They are now
CC produced free of PDGF A, dimers and heterodimers. The antibodies are
CC useful for affinity purification.
CC See also P80595-96, P80597-98 and N81202-03.
SQ Sequence 241 AA;

Query Match 5.9%; Score 6; DB 1; Length 241;
Best Local Similarity 100.0%; Pred. No. 7.69e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 95 AECKTR 100
||||||
Qy 71 AECKTR 76

RESULT 59
R63469 standard; protein; 241 AA.
AC R63469;
DT 28-JUN-1995 (first entry)
DE Recombinant platelet derived growth factor B c-sis.
KW Recombinant platelet derived growth factor B c-sis; rPDGF B c-sis;
KW wound healing; antibody production.
OS Synthetic.
PN EP-622456-A.
PD 02-NOV-1994.
PF 10-MAR-1988; 302116.
PR 13-MAR-1987; US-025344.
PR 19-FEB-1988; US-152045.
PA (AMGE-) AMGEN INC.
PI Nicolson MA, Thomason AR;
DR WPI; 94-334642/42.
PT New vectors expressing platelet-derived growth factor B
PT epitope(s) - and contg. c-sis or v-sis genes portions; are used
PT to promote wound healing and produce antibodies
PS Disclosure; Fig 2; 49pp; English.
CC R63469 describes the amino acid sequence of recombinant platelet
CC derived growth factor B c-sis (rPDGF B c-sis). It may be used to
CC promote wound healing, or as an antigen to prepare antisera or
CC monoclonal antibodies specific for epitopes on the B chain of PDGF.
SQ Sequence 241 AA;

Query Match 5.9%; Score 6; DB 1; Length 241;
Best Local Similarity 100.0%; Pred. No. 7.69e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 95 AECKTR 100
||||||
Qy 71 AECKTR 76

RESULT 60
ID W31729 standard; Protein; 261 AA.
AC W31729;
DT 14-APR-1998 (first entry)
DE Mycobacterium tuberculosis sigma factor.
KW Sigma factor; sigF gene; tuberculosis; latency; diagnosis;
KW therapy; vaccine.
OS Mycobacterium tuberculosis H37Rv.
FH Key Location/Qualifiers
FT Misc_difference 214
FT /note= "encoded by SCC"
PN WO9735611-A1.
PD 02-OCT-1997.
PF 27-MAR-1997; U03457.
PR 27-MAR-1996; US-622353.
PR 27-MAR-1996; US-622352.
PA (UYJO) UNIV JOHNS HOPKINS.
PI Bishai WR, Demaio J, Young DB, Zhang Y;
DR WPI; 97-489391/45.
DR N-PSDB; T88842.
PT DNA encoding Mycobacterium tuberculosis sigma factor or ORF X or Y
PT - used to detect presence of latent pathogenic mycobacterium and to
PT identify compounds that regulate binding of M. tuberculosis sigF to
PT orfX
PS Claim 7; Page 51-52; 72pp; English.
CC This protein comprises the stress response sigma factor (sigF) of
CC Mycobacterium tuberculosis. Its amino acid sequence was determined
CC from an isolated genomic DNA molecule (see T88842) of M.
CC tuberculosis H37Rv. 2 Genes, orfX and orfY, regulate sigF
CC activity. The presence of a latent pathogenic mycobacterium in a
CC human can be determined by detecting M. tuberculosis sigF nucleic
CC acid or protein in a body sample (claimed). A cell containing a
CC sigF reporter construct can be used to screen for potential
CC therapeutic agents that trigger the growth arrest of M.
CC tuberculosis by activating the expression of sigF, or reactivate
CC latent M. tuberculosis by inhibiting the expression of sigF
CC (claimed). Measuring in vitro transcription from a transcription
CC construct incubated with M. tuberculosis in the presence or absence
CC of a test compound can be used to determine if the compound is a
CC potential therapeutic agent for use in modulating the growth of M.
CC tuberculosis by regulating the activity of M. tuberculosis sigF
CC (claimed). The M. tuberculosis sigF, orfX and orfY proteins (see
CC W31729-31) can be used to identify compounds which regulate the
CC binding of M. tuberculosis sigF to orfX. A claimed tuberculosis
CC vaccine comprises an M. tuberculosis strain that has a mutation
CC which disrupts sigF expression or translation.
SQ Sequence 261 AA;

Query Match 5.9%; Score 6; DB 1; Length 261;
Best Local Similarity 100.0%; Pred. No. 7.69e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 145 SELAAE 150
||||||
Qy 67 SELAAE 72

RESULT 61
ID W89843 standard; Protein; 264 AA.
AC W89843;
DT 18-FEB-1999 (first entry)
DE Protein encoded by clone g9 ORF3.
KW Antigen; immunogenic cluster family; vaccine; gastritis; diagnosis;
KW peptic ulcer; gastric adenocarcinoma; gastric lymphoma.
OS Helicobacter pylori.
PN WO9849314-A2.
PD 05-NOV-1998.
PF 27-APR-1998; U08487.
PR 14-OCT-1997; US-061958.
PR 25-APR-1997; US-045107.
PA (GENE-) GENELABS TECHNOLOGIES INC.
PI Chow TP, Fry KE, Lim MY, McAtee CP;
DR WPI; 99-009433/01.
DR N-PSDB; V90772.
PT New Helicobacter pylori antigens and related nucleic acid sequences
PT - useful in serological diagnosis and protective vaccines, providing
PT long-lasting immune response
PS Claim 15; Page 190-191; 402pp; English.
CC The present sequence represents a Helicobacter pylori antigenic protein
CC that is characterised by immunoreactivity with H. pylori-positive
CC antisera. The proteins are highly immunogenic and induce a long-lasting
CC immune response that persists even after antimicriobial treatment. In
CC antibody-detection assays, on sera, plasma, urine, saliva etc., they are
CC highly sensitive and specific. The specification also describes 69
CC previously unrecognised immunogenic cluster families. H. pylori antigens
CC are used to detect H. pylori-specific antibodies, for diagnosing
CC infection or to confirm eradication of infection, and in vaccines to
CC protect against H. pylori infection and related diseases (gastritis,
CC peptic ulcer, gastric adenocarcinoma/lymphoma).
SQ Sequence 264 AA;

```
  Query Match           5.9%;  Score 6;  DB 1;  Length 264;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches    6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      88 SVLTST 93
           ||||||
Qy      16 SVLTST 21


RESULT   62
ID   W89920 standard; Protein; 265 AA.
AC   W89920;
DT   18-FEB-1999  (first entry)
DE   Antigen 5 from cluster 38e.
KW   Antigen; immunogenic cluster family; vaccine; gastritis; diagnosis;
KW   peptic ulcer; gastric adenocarcinoma; gastric lymphoma.
OS   Helicobacter pylori.
PN   WO9849314-A2.
PD   05-NOV-1998.
PF   27-APR-1998; U08487.
PR   14-OCT-1997; US-061958.
PR   25-APR-1997; US-045107.
PA   (GENE-) GENELABS TECHNOLOGIES INC.
PI   Chow TP, Fry KE, Lim MY, McAtee CP;
DR   WPI; 99-009433/01.
PT   New Helicobacter pylori antigens and related nucleic acid sequences
PT   - useful in serological diagnosis and protective vaccines, providing
PT   long-lasting immune response
PS   Claim 1; Page 250; 402pp; English.
CC   The present sequence represents a Helicobacter pylori antigenic protein
CC   that is characterised by immunoreactivity with H. pylori-positive
CC   antisera. The proteins are highly immunogenic and induce a long-lasting
CC   immune response that persists even after antimicriobial treatment. In
CC   antibody-detection assays, on sera, plasma, urine, saliva etc., they are
CC   highly sensitive and specific. The specification also describes 69
CC   previously unrecognised immunogenic cluster families. H. pylori antigens
CC   are used to detect H. pylori-specific antibodies, for diagnosing
CC   infection or to confirm eradication of infection, and in vaccines to
CC   protect against H. pylori infection and related diseases (gastritis,
CC   peptic ulcer, gastric adenocarcinoma/lymphoma).
SQ   Sequence   265 AA;

  Query Match           5.9%;  Score 6;  DB 1;  Length 265;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches    6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      88 SVLTST 93
           ||||||
Qy      16 SVLTST 21


RESULT   63
ID   R40963 standard; protein; 271 AA.
AC   R40963;
DT   25-FEB-1994  (first entry)
DE   PDGF Bv-sis.
KW   Monoclonal antibody; specific; epitope; B-Chain; PDGF;
KW   platelet-derived growth factor; wound healing; binding.
OS   Simian sarcoma virus.
PN   EP-559234-A.
PD   08-SEP-1993.
PF   10-MAR-1988; 302116.
PR   13-MAR-1987; US-025344.
PR   19-FEB-1988; US-152045.
PA   (AMGE-) AMGEN.
PI   Nicolson MA,  Thomason AR;
DR   WPI; 93-282094/36.
PT   New monoclonal antibody for prepn. of poly:peptide for wound
PT   healing - is specific for epitope in B-chain of platelet-derived
PT   growth factor
PS   Example 1; Fig 1; 39pp; English.
CC   Example 1 describes the subcloning and analysis of C-sis and V-sis
CC   genes encoding rPDGF B chain. Monoclonal antibodies against
CC   rPDGF Bv-sis have the affinity to specifically bind to an epitope
CC   found in the B chain of PDGF.
SQ   Sequence   271 AA;

  Query Match           5.9%;  Score 6;  DB 1;  Length 271;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches    6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     125 AECKTR 130
           ||||||
Qy      71 AECKTR 76


RESULT   64
ID   P80595 standard; protein; 271 AA.
AC   P80595;
DT   12-OCT-1990  (first entry)
DE   Recombinant platelet derived growth factor Bv-sis.
KW   REcombinant platelet derived growth factor; wound healing;
KW   fusion peptides;  Bv-sis; antibodies.
OS   Simian sarcoma virus.
PN   EP-282317-A.
PD   14-SEP-1988.
PF   10-MAR-1988; 302116.
PR   19-FEB-1988; US-152045.
PA   (AMGE-) AMGEN INC.
PI   Thomason AR, Nicolson; MA;
DR   WPI; 88-259036/37.
PT   Purified recombinant platelet derived growth factor -
PT   for accelerating wound healing, and fusion polypeptide(s),
PT   encoding DNA sequences and derived antibodies.
PS   Disclosure; p; English.
CC   The peptide has a sufficient part of the structural conformation of
CC   PDGF B, an epitope for binding to a monoclonal antibody spaecific
CC   for an epitope of a B chain PDGF, one or more of the biological
CC   properties of naturally occurring PDGF and it has purity greater than
CC   95%. The cpds. is used in veterinary and human medicine for topical
CC   application to wounds to accelerate healing. They are produced free of
CC   PDGF A, dimers and heterodimers. The antibodies are useful for
CC   affinity purification. See also P80595-P80598 and N81202-N81203.
SQ   Sequence   271 AA;

  Query Match           5.9%;  Score 6;  DB 1;  Length 271;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches    6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     125 AECKTR 130
           ||||||
Qy      71 AECKTR 76


RESULT   65
ID   R63468 standard; protein; 271 AA.
AC   R63468;
DT   28-JUN-1995  (first entry)
DE   Recombinant platelet derived growth factor B v-sis.
KW   Recombinant platelet derived growth factor B v-sis; rPDGF B v-sis;
KW   wound healing; antibody production.
OS   Synthetic.
PN   EP-622456-A.
PD   02-NOV-1994.
PF   10-MAR-1988; 302116.
PR   13-MAR-1987; US-025344.
PR   19-FEB-1988; US-152045.
PA   (AMGE-) AMGEN INC.
PI   Nicolson MA,  Thomason AR;
DR   WPI; 94-334642/42.
PT   New vectors expressing platelet-derived growth factor B
PT   epitope(s)  - and contg. c-sis or v-sis genes portions; are used
PT   to promote wound healing and produce antibodies
PS   Disclosure; Fig 1; 49pp; English.
CC   R63468 describes the amino acid sequence of recombinant platelet
CC   derived growth factor B v-sis (rPDGF B v-sis). It may be used to
```

```
CC   promote wound healing, or as an antigen to prepare antisera or
CC   monoclonal antibodies specific for epitopes on the B chain of PDGF.
SQ   Sequence   271 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 271;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    125 AECKTR 130
          ||||||
Qy     71 AECKTR 76


RESULT   66
ID   R60616 standard; Protein; 282 AA.
AC   R60616;
DT   22-MAY-1995  (first entry)
DE   Human PDGF-B 119 linked to PDGF-B 109 via pre-pro region linker.
KW   Human platelet derived growth factor; PDGF-B subunit; v-sis gene;
KW   PDGF-BB fusion protein; fusion dimer; protein refolding;
K    wound healing.
O    Homo sapiens.
FH   Key             Location/Qualifiers
FT   protein         1. .119
FT                   /label= PDGF-B_119_subunit
FT   region          120. .176
FT                   /label= linker
FT                   /note= "amino acids -54 to -1 of the pre-pro region
FT                   of the PDGF-B precursor protein"
FT   protein         177. .282
FT                   /label= PDGF-B_109_subunit
PN   EP-618227-A.
PD   05-OCT-1994.
PF   31-MAR-1994; 105075.
PR   01-APR-1993; US-041635.
PA   (AMGE-) AMGEN INC.
PI   Thomason AR;
DR   WPI; 94-304405/38.
PT   New multimeric polypeptide comprising fused subunits of natural
PT   protein.  - and related DNA and transformed cells, esp. dimers of
PT   platelet derived growth factor, useful for stimulating healing of
PT   wounds.
PS   Claim 12; Fig 1; 30pp; English.
CC   The PDGF-B 109 subunit coding sequence was obtained as composite
CC   sequence of fragments from human PDGF-B pre-pro region, human v-sis
CC   gene and synthetic fragments. The mature 109 amino acid subunit was
CC   incorporated into a single polypeptide with a PDGF-B 119 subunit,
CC   separated by a peptide linker to produce the preferred fusion
C    polypeptide of the invention (R60616). The fusion dimer is more
C    easily and rapidly refolded than unfused multimers and is
CC   useful for treating wounds.
SQ   Sequence   282 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 282;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     14 AECKTR 19
          ||||||
Qy     71 AECKTR 76


RESULT   67
ID   W04936 standard; Protein; 293 AA.
AC   W04936;
DT   20-MAY-1997  (first entry)
DE   Chimeric endoglucanase #4.
KW   Cellulytic enzyme; endoglucanase; hydrolysis; cellulose; microorganism;
KW   plant cellulase; catalytic region; textile; backstaining; bio-polishing;
KW   stone-washing; cellulosic fabric; colour clarification; defibration;
KW   cell wall degradation; paper pulp; debarking; fibre modification;
KW   enzymatic de-inking; drainage improvement; chimera.
OS   Chimera: Crinipellis scabella.
OS   Chimera: Humicoal insolens.
FH   Key             Location/Qualifiers
FT   region          1. .221
FT                   /note= "C.scabella endoglucanase"
FT   region          222. .293
FT                   /note= "C-terminus of H.insolens endoglucanase"
PN   WO9629397-A1.
PD   26-SEP-1996.
PF   18-MAR-1996; DK0105.
PR   17-MAR-1995; DK-000272.
PR   08-AUG-1995; DK-000887.
PR   08-AUG-1995; DK-000885.
PR   08-AUG-1995; DK-000888.
PR   08-AUG-1995; DK-000886.
PR   12-FEB-1996; DK-000137.
PA   (NOVO ) NOVO-NORDISK AS.
PI   Andersen LN,  Ihara M,  Kauppinen MS,  Lange L,  Lassen SF;
PI   Nielsen RI, Schuelein M,  Takagi S;
DR   WPI; 96-443173/44.
DR   N-PSDB; T39074.
PT   New endo:glucanase enzyme preparations - contg. conserved catalytic
PT   regions, useful for treating fabrics, textiles, plant material or
PT   paper pulp
PS   Example 4; Page 176-178; 316pp; English.
CC   This sequence represents a chimeric enzyme, containing an enzyme of the
CC   invention (see W04930 for full length C. scabella endoglucanase).  The
CC   enzymes of the invention possess cellulytic (particularly endoglucanase)
CC   activity.  Cellulytic enzymes are involved in the the hydrolysis of
CC   cellulose, and are synthesised by a large number of microorganisms and
CC   plants.  The enzymes of the invention containing the conserved catalytic
CC   regions (such as W04913) exhibit improved performance, e.g. 50 times
CC   higher performance, compared to multiple domain enzymes.  The enzymes can
CC   be used for the treatment of fabrics or textiles, preferably for
CC   preventing backstaining, for bio-polishing or for stone-washing
CC   cellulosic fabric. They can also be used to provide colour clarification
CC   for laundry. The enzymes can also be used for the degradation or
CC   modification of plant material, such as cell walls.  They can also be
CC   used in the treatment of paper pulp preferably for debarking,
CC   defibration, fibre modification, enzymatic de-inking or drainage
CC   improvement.
SQ   Sequence   293 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 293;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     94 LAGSSE 99
          ||||||
Qy     63 LAGSSE 68


RESULT   68
ID   W04933 standard; Protein; 297 AA.
AC   W04933;
DT   20-MAY-1997  (first entry)
DE   Chimeric endoglucanase #1.
KW   Cellulytic enzyme; endoglucanase; hydrolysis; cellulose; microorganism;
KW   plant cellulase; catalytic region; textile; backstaining; bio-polishing;
KW   stone-washing; cellulosic fabric; colour clarification; defibration;
KW   cell wall degradation; paper pulp; debarking; fibre modification;
KW   enzymatic de-inking; drainage improvement; chimera.
OS   Chimera: Myceliophthora thermophila.
OS   Chimera: Humicoal insolens.
FH   Key             Location/Qualifiers
FT   region          1. .225
FT                   /note= "M.thermophila endoglucanase"
FT   region          226. .297
FT                   /note= "C-terminus of H.insolens endoglucanase"
PN   WO9629397-A1.
PD   26-SEP-1996.
PF   18-MAR-1996; DK0105.
PR   17-MAR-1995; DK-000272.
PR   08-AUG-1995; DK-000887.
```

PR 08-AUG-1995; DK-000885.
PR 08-AUG-1995; DK-000888.
PR 08-AUG-1995; DK-000886.
PR 12-FEB-1996; DK-000137.
PA (NOVO) NOVO-NORDISK AS.
PI Andersen LN, Ihara M, Kauppinen MS, Lange L, Lassen SF;
PI Nielsen RI, Schuelein M, Takagi S;
DR WPI; 96-443173/44.
DR N-PSDB; T39061.
PT New endo:glucanase enzyme preparations - contg. conserved catalytic
PT regions, useful for treating fabrics, textiles, plant material or
PT paper pulp
PS Example 1; Page 144-146; 316pp; English.
CC This sequence represents a chimeric enzyme, containing an enzyme of the
CC invention (see W04925 for full length M. thermophila endoglucanase). The
CC enzymes of the invention possess cellulytic (particularly endoglucanase)
CC activity. Cellulytic enzymes are involved in the the hydrolysis of
CC cellulose, and are synthesised by a large number of microorganisms and
CC plants. The enzymes of the invention containing the conserved catalytic
CC regions (such as W04913) exhibit improved performance, e.g. 50 times
CC higher performance, compared to multiple domain enzymes. The enzymes can
CC be used for the treatment of fabrics or textiles, preferably for
CC preventing backstaining, for bio-polishing or for stone-washing
CC cellulosic fabric. They can also be used to provide colour clarification
CC for laundry. The enzymes can also be used for the degradation or
CC modification of plant material, such as cell walls. They can also be
CC used in the treatment of paper pulp preferably for debarking,
CC defibration, fibre modification, enzymatic de-inking or drainage
CC improvement.
SQ Sequence 297 AA;

Query Match 5.9%; Score 6; DB 1; Length 297;
Best Local Similarity 100.0%; Pred. No. 7.69e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 98 LAGSSE 103
|||||
Qy 63 LAGSSE 68

RESULT 69
ID W04935 standard; Protein; 298 AA.
AC W04935;
DT 20-MAY-1997 (first entry)
DE Chimeric endoglucanase #3.
KW Cellulytic enzyme; endoglucanase; hydrolysis; cellulose; microorganism;
KW plant cellulase; catalytic region; textile; backstaining; bio-polishing;
KW stone-washing; cellulosic fabric; colour clarification; defibration;
KW cell wall degradation; paper pulp; debarking; fibre modification;
KW enzymatic de-inking; drainage improvement; chimera.
OS Chimera: Crinipellis scabella.
OS Chimera: Humicoal insolens.
FH Key Location/Qualifiers
FT region 1. .226
FT /note= "C.scabella endoglucanase"
FT region 227. .298
FT /note= "C-terminus of H.insolens endoglucanase"
PN WO9629397-A1.
PD 26-SEP-1996.
PF 18-MAR-1996; DK0105.
PR 17-MAR-1995; DK-000272.
PR 08-AUG-1995; DK-000887.
PR 08-AUG-1995; DK-000885.
PR 08-AUG-1995; DK-000888.
PR 08-AUG-1995; DK-000886.
PR 12-FEB-1996; DK-000137.
PA (NOVO) NOVO-NORDISK AS.
PI Andersen LN, Ihara M, Kauppinen MS, Lange L, Lassen SF;
PI Nielsen RI, Schuelein M, Takagi S;
DR WPI; 96-443173/44.
DR N-PSDB; T39073.
PT New endo:glucanase enzyme preparations - contg. conserved catalytic
PT regions, useful for treating fabrics, textiles, plant material or
PT paper pulp
PS Example 4; Page 172-174; 316pp; English.
CC This sequence represents a chimeric enzyme, containing an enzyme of the
CC invention (see W04930 for full length C. scabella endoglucanase). The
CC enzymes of the invention possess cellulytic (particularly endoglucanase)
CC activity. Cellulytic enzymes are involved in the the hydrolysis of
CC cellulose, and are synthesised by a large number of microorganisms and
CC plants. The enzymes of the invention containing the conserved catalytic
CC regions (such as W04913) exhibit improved performance, e.g. 50 times
CC higher performance, compared to multiple domain enzymes. The enzymes can
CC be used for the treatment of fabrics or textiles, preferably for
CC preventing backstaining, for bio-polishing or for stone-washing
CC cellulosic fabric. They can also be used to provide colour clarification
CC for laundry. The enzymes can also be used for the degradation or
CC modification of plant material, such as cell walls. They can also be
CC used in the treatment of paper pulp preferably for debarking,
CC defibration, fibre modification, enzymatic de-inking or drainage
CC improvement.
SQ Sequence 298 AA;

Query Match 5.9%; Score 6; DB 1; Length 298;
Best Local Similarity 100.0%; Pred. No. 7.69e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 99 LAGSSE 104
|||||
Qy 63 LAGSSE 68

RESULT 70
ID W81823 standard; Protein; 307 AA.
AC W81823;
DT 02-FEB-1999 (first entry)
DE Human VTAP-2 protein.
KW VTAP-2; human; vesicle transport associated protein; treatment; human;
KW disorder; cancer; reproduction; immune; developmental.
OS Homo sapiens.
PN US5840539-A.
PD 24-NOV-1998.
PF 10-OCT-1997; 948616.
PR 10-OCT-1997; US-948616.
PA (INCY-) INCYTE PHARM INC.
PI Corley NC, Hillman JL, Lal P, Shah P;
DR WPI; 99-034038/03.
DR N-PSDB; V64837.
PT DNA encoding vesicle transport-associated protein - useful for
PT producing recombinant protein
PS Example VIII; Fig 2A-D; 48pp; English.
CC This sequence represents a novel human vesicle transport associated
CC protein (VSAP-2). This protein could be used in methods for treating
CC disorders associated with the expression of VTAP, such as cancer,
CC immune, reproductive and developmental disorders.
SQ Sequence 307 AA;

Query Match 5.9%; Score 6; DB 1; Length 307;
Best Local Similarity 100.0%; Pred. No. 7.69e+01;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 133 STKINL 138
|||||
Qy 82 STKINL 87

RESULT 71
ID W04934 standard; Protein; 308 AA.
AC W04934;
DT 20-MAY-1997 (first entry)
DE Chimeric endoglucanase #2.
KW Cellulytic enzyme; endoglucanase; hydrolysis; cellulose; microorganism;
KW plant cellulase; catalytic region; textile; backstaining; bio-polishing;
KW stone-washing; cellulosic fabric; colour clarification; defibration;
KW cell wall degradation; paper pulp; debarking; fibre modification;
KW enzymatic de-inking; drainage improvement; chimera.

```
OS   Chimera: Myceliophthora thermophila.
OS   Chimera: Humicoal insolens.
FH   Key             Location/Qualifiers
FT   region          1. .225
FT                   /note= "M.thermophila endoglucanase"
FT   region          226. .308
FT                   /note= "C-terminus of H.insolens endoglucanase"
PN   WO9629397-A1.
PD   26-SEP-1996.
PF   18-MAR-1996; DK0105.
PR   17-MAR-1995; DK-000272.
PR   08-AUG-1995; DK-000887.
PR   08-AUG-1995; DK-000885.
PR   08-AUG-1995; DK-000888.
PR   08-AUG-1995; DK-000886.
PR   12-FEB-1996; DK-000137.
PA   (NOVO ) NOVO-NORDISK AS.
PI   Andersen LN,  Ihara M,  Kauppinen MS,  Lange L,  Lassen SF;
PI   Nielsen RI, Schuelein M,  Takagi S;
DR   WPI; 96-443173/44.
DR   N-PSDB; T39062.
PT   New endo:glucanase enzyme preparations - contg. conserved catalytic
PT   regions, useful for treating fabrics, textiles, plant material or
PT   paper pulp
PS   Example 1; Page 148-150; 316pp; English.
CC   This sequence represents a chimeric enzyme, containing an enzyme of the
CC   invention (see W04925 for full length M. thermophila endoglucanase).  The
CC   enzymes of the invention possess cellulytic (particularly endoglucanase)
CC   activity.  Cellulytic enzymes are involved in the the hydrolysis of
CC   cellulose, and are synthesised by a large number of microorganisms and
CC   plants.  The enzymes of the invention containing the conserved catalytic
CC   regions (such as W04913) exhibit improved performance, e.g. 50 times
CC   higher performance, compared to multiple domain enzymes.  The enzymes can
CC   be used for the treatment of fabrics or textiles, preferably for
CC   preventing backstaining, for bio-polishing or for stone-washing
CC   cellulosic fabric. They can also be used to provide colour clarification
CC   for laundry. The enzymes can also be used for the degradation or
CC   modification of plant material, such as cell walls.  They can also be
CC   used in the treatment of paper pulp preferably for debarking,
CC   defibration, fibre modification, enzymatic de-inking or drainage
CC   improvement.
SQ   Sequence   308 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 308;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db      98 LAGSSE 103
           ||||||
Qy      63 LAGSSE 68


RESULT   72
ID   W27267 standard; Protein; 348 AA.
AC   W27267;
DT   28-APR-1998  (first entry)
DE   Streptococcus pneumoniae glutamyl tRNA synthetase.
KW   Glutamyl tRNA synthetase; gluS protein; bacterial infection; vaccine;
KW   meningitis; antibacterial; immune response.
OS   Streptococcus pneumoniae.
PN   WO9738718-A1.
PD   23-OCT-1997.
PF   18-APR-1997; U06753.
PR   18-APR-1996; GB-007992.
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
PI   Lawlor EJ;
DR   WPI; 97-526211/48.
DR   N-PSDB; T91268.
PT   New isolated nucleic acid encoding glutamyl tRNA synthetase of
PT   Streptococcus pneumoniae - useful for diagnosis, treatment and
PT   prevention of bacterial infections, especially meningitis
PS   Claim 12; Page 35-36; 42pp; English.
CC   The present sequence represents glutamyl tRNA synthetase (gluS) from
CC   Streptococcus pneumoniae 0100993 (NCIMB 40794). The polynucleotide
CC   encoding gluS is used to express recombinant gluS, i.e. the gluS
CC   polypeptide or their fragments, which are used to treat conditions that
CC   require gluS activity, also as antisense sequences to control expression
CC   of gluS. GluS, or vectors that express it, are used to induce an immune
CC   (antibody and/or T cell) response, specifically for protection against
CC   S. pneumoniae infection or to screen for (ant)agonists of the
CC   polynucleotide/protein activity, particularly antibacterials.
CC   Antagonists, e.g. Ab, are used to treat conditions requiring inhibition
CC   of gluS, generally any S. pneumoniae infection but particularly
CC   meningitis. Fragments of the polynucleotide encoding gluS are useful as
CC   probes to isolate full-length or related sequences, or for diagnosis,
CC   e.g. by polymerase chain reaction, of the stage and type of an
CC   infection, including detection of mutations and polymorphisms. Diagnosis
CC   may also be done by detecting overexpression of the gluS genes, e.g. by
CC   immunoassay. Ab are used to treat infections; to isolate/identify gluS-
CC   expressing clones; to purify gluS and as immunoassay reagents. More
CC   generally, gluS encoding polynucleotide-antagonists can prevent adhesion
CC   of bacteria to wounds, in-dwelling devices; block gluS-protein mediated
CC   invasion of mammalian cells and block normal progression of infection.
SQ   Sequence   348 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 348;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     212 DDHIAN 217
           ||||||
Qy      88 DDHIAN 93


RESULT   73
ID   R98154 standard; Protein; 355 AA.
AC   R98154;
DT   06-JAN-1997  (first entry)
DE   Blowfly PM48 antigen.
KW   Blowfly strike; cutaneous myiasis; PM48; antigen; vaccine.
OS   Lucilia cuprina.
FH   Key             Location/Qualifiers
FT   modified_site   97. .99
FT                   /label= Glycosylation
FT                   /note= "potential N-linked glycosylation site"
FT   modified_site   340. .342
FT                   /label= Glycosylation
FT                   /note= "potential N-linked glycosylation site"
PN   AU9531744-A.
PD   18-APR-1996.
PF   20-SEP-1995; 031744.
PR   29-SEP-1994; AU-008452.
PA   (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PI   Eisemann CH,  Schorderet SA,  Tellam RL;
DR   WPI; 96-222457/23.
DR   N-PSDB; T30074.
PT   PM48 antigen from Lucilia cuprina and related DNA - useful in
PT   vaccines to protect sheep against blowfly strike
PS   Claim 1; Fig 5; 50pp; English.
CC   PM48 antigen (R98154) is a Lucilia cuprina peritrophic membrane
CC   protein.  Its amino acid sequence was deduced from a cDNA clone
CC   (T30074) obtd. from first instar larvae.  Large quantities of
CC   recombinant PM48 antigen can be produced in transformed host cells.
CC   Vaccines comprising the PM48 antigen are useful for the prophylaxis
CC   or treatment of blowfly strike in sheep.  The vaccine is effective
CC   against blowfly strike due to blowflies from genera including
CC   Lucilia, Calliphora, Chrysomya and Cochliomyia.
SQ   Sequence   355 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 355;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db     188 CERTNG 193
           ||||||
```

```
Qy     22 CERTNG 27


RESULT   74
ID   W79743 standard; Protein; 374 AA.
AC   W79743;
DT   02-FEB-1999  (first entry)
DE   Crepis sp. delta-12-epoxygenase.
KW   Fatty acid epoxygenase; CrepX gene; delta-12-epoxygenase;
KW   mixed function monooxygenase; epoxygenated fatty acid;
KW   transgenic plant; vegetable oil; oilseed.
OS   Crepis sp.
FH   Key             Location/Qualifiers
FT   Region          99. .104
FT                   /note= "His-rich region"
FT   Region          135. .139
FT                   /note= "His-rich region"
FT   Region          309. .313
FT                   /note= "His-rich region"
F    Misc_difference 293
I                    /note= "encoded by AGN"
FT   Misc_difference 305
FT                   /note= "encoded by GTN"
PN   WO9846762-A1.
PD   22-OCT-1998.
PF   09-APR-1998; AU0246.
PR   20-JUN-1997; US-050403.
PR   15-APR-1997; AU-006223.
PR   15-APR-1997; AU-006226.
PR   16-APR-1997; US-043706.
PA   (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA   (STYM/) STYMNE S.
PI   Green A, Lenman M, Singh S, Stymne S;
DR   WPI; 98-568734/48.
DR   N-PSDB; V63102.
PT   New isolated fatty acid epoxygenase gene - used particularly for
PT   transforming plants for producing modified oils for use in, e.g.
PT   coatings, resins, glues, plastics, surfactants or lubricants
PS   Claim 29; Page 84-86; 150pp; English.
CC   This is the amino acid sequence of a novel epoxygenase of a
CC   vernolic acid-containing Crepis sp. (not crepis palaestina).  It
CC   was deduced from isolated cDNA clone CrepX (see V63102).  The
CC   deduced sequence contains His-rich motifs (see W79752-54) that are
CC   characteristic of mixed function monooxygenases, and shows a high
CC   degree of homology to a novel delta-12-epoxygenase (see W79742) of
CC   C. palaestina.  The invention relates generally to novel genetic
CC   sequences (see V63101-03) encoding fatty acid epoxygenases (see
C    W79742-44), especially delta-12-epoxygenases or mixed function
     monooxygenases.  These provide the means by which fatty acid
C    metabolism can be manipulated in e.g. yeast, mould, bacteria,
CC   insects, birds, mammals and plants (especially oilseed plants such
CC   as flax), in particular to convert unsaturated fatty acids to
CC   epoxygenated fatty acids.  The invention includes genetically
CC   modified oil-accumulating organisms and to the oils derived from
CC   them.  These oils can be used in production of coatings, resins,
CC   glues, plastics, surfactants or lubricants.
SQ   Sequence   374 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 374;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    141 NTSSID 146
          ||||||
Qy     30 NTSSID 35


RESULT   75
ID   W79742 standard; Protein; 374 AA.
AC   W79742;
DT   02-FEB-1999  (first entry)
DE   Crepis palaestina delta-12-epoxygenase.
KW   Fatty acid epoxygenase; Cpal2 gene; mixed function monooxygenase;
KW   delta-12-epoxygenase; epoxygenated fatty acid; transgenic plant;
KW   vegetable oil; oilseed.
OS   Crepis palaestina.
FH   Key             Location/Qualifiers
FT   Region          99. .104
FT                   /note= "His-rich region"
FT   Region          135. .139
FT                   /note= "His-rich region"
FT   Region          309. .313
FT                   /note= "His-rich region"
PN   WO9846762-A1.
PD   22-OCT-1998.
PF   09-APR-1998; AU0246.
PR   20-JUN-1997; US-050403.
PR   15-APR-1997; AU-006223.
PR   15-APR-1997; AU-006226.
PR   16-APR-1997; US-043706.
PA   (CSIR ) COMMONWEALTH SCI & IND RES ORG.
PA   (STYM/) STYMNE S.
PI   Green A, Lenman M, Singh S, Stymne S;
DR   WPI; 98-568734/48.
DR   N-PSDB; V63101.
PT   New isolated fatty acid epoxygenase gene - used particularly for
PT   transforming plants for producing modified oils for use in, e.g.
PT   coatings, resins, glues, plastics, surfactants or lubricants
PS   Claim 29; Page 78-81; 150pp; English.
CC   This is the amino acid sequence of a novel mixed function
CC   monoxygenase of Crepis palaestina that is characterised as having
CC   delta-12-epoxygenase activity.  It was deduced from isolated
CC   full-length cDNA clone Cpal2 (see V63101).  The deduced sequence
CC   contains His-rich motifs (see W79752-54) that are characteristic of
CC   delta-12-desaturase enzymes.  The Cpal2 gene was shown to be highly
CC   expressed in developing seeds, with no expression detectable in
CC   leaves.  The invention relates generally to novel genetic sequences
CC   (see V63101-03) encoding fatty acid epoxygenases (see W79742-44),
CC   especially delta-12-epoxygenases or mixed function monooxygenases.
CC   These provide the means by which fatty acid metabolism can be
CC   manipulated in e.g. yeast, mould, bacteria, insects, birds, mammals
CC   and plants (especially oilseed plants such as flax), in particular
CC   to convert unsaturated fatty acids to epoxygenated fatty acids.
CC   The invention extends to genetically modified oil-accumulating
CC   organisms and to the oils derived from them.  These oils can be
CC   used in production of coatings, resins, glues, plastics,
CC   surfactants or lubricants.
SQ   Sequence   374 AA;

  Query Match             5.9%;  Score 6;  DB 1;  Length 374;
  Best Local Similarity 100.0%;  Pred. No. 7.69e+01;
  Matches     6;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Db    141 NTSSID 146
          ||||||
Qy     30 NTSSID 35


Search completed: Wed Aug 16 09:59:07 2000
Job time : 43 secs.
```